

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:34:35 ; Search time 30 Seconds  
(without alignments)  
746.837 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNYKRFVVGVTSTFVFFLS.....ASATLNTDYFGWEIGRPFAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5 COMB.pdp.\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB.pdp.\*  
3: /cgn2\_6/ptodata/1/iaa/H COMB.pdp.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pdp.\*  
5: /cgn2\_6/ptodata/1/iaa/RE COMB.pdp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	271	2	US-09-660-587-46
2	1434	100.0	271	2	US-09-811-007A-46
3	509.5	35.5	280	2	US-09-660-587-42
4	509.5	35.5	280	2	US-09-314-701-48
5	509.5	35.5	280	2	US-09-811-007A-42
6	509.5	35.5	280	2	US-10-314-639-48
7	509.5	35.5	280	2	US-10-059-964A-48
8	507	35.4	283	2	US-09-660-587-10
9	507	35.4	283	2	US-09-261-358A-10
10	507	35.4	283	2	US-09-201-458-6
11	507	35.4	283	2	US-09-314-701-4
12	507	35.4	283	2	US-09-811-007A-10
13	507	35.4	283	2	US-10-314-639-4
14	507	35.4	283	2	US-10-059-964A-4
15	492	34.3	285	2	US-09-314-701-30
16	492	34.3	285	2	US-10-314-639-30
17	492	34.3	285	2	US-10-059-964A-30
18	489	34.1	279	2	US-09-314-701-24
19	489	34.1	279	2	US-10-314-639-24
20	489	34.1	279	2	US-10-059-964A-24
21	463	32.3	283	2	US-09-314-701-26
22	463	32.3	283	2	US-10-314-639-26
23	463	32.3	283	2	US-10-059-964A-26
24	455.5	31.8	298	2	US-09-314-701-52
25	455.5	31.8	298	2	US-10-314-639-52
26	455.5	31.8	298	2	US-10-059-964A-52
27	455	31.7	281	2	US-09-314-701-58

NO sum

28	455	31.7	281	2	US-10-314-639-58	Sequence 58, Appl
29	455	31.7	281	2	US-10-059-964A-58	Sequence 58, Appl
30	440	30.7	291	2	US-09-314-701-18	Sequence 18, Appl
31	440	30.7	291	2	US-10-314-639-18	Sequence 18, Appl
32	440	30.7	291	2	US-10-059-964A-18	Sequence 18, Appl
33	437	30.5	293	2	US-09-660-587-40	Sequence 40, Appl
34	437	30.5	293	2	US-09-811-007A-40	Sequence 40, Appl
35	435.5	30.4	288	2	US-09-314-701-32	Sequence 32, Appl
36	435.5	30.4	288	2	US-10-314-639-32	Sequence 32, Appl
37	435.5	30.4	288	2	US-10-059-964A-32	Sequence 32, Appl
38	432	30.1	293	2	US-09-314-701-44	Sequence 44, Appl
39	432	30.1	293	2	US-10-314-639-44	Sequence 44, Appl
40	432	30.1	293	2	US-10-059-964A-44	Sequence 44, Appl
41	431.5	30.1	300	2	US-09-314-701-50	Sequence 50, Appl
42	431.5	30.1	300	2	US-10-314-639-50	Sequence 50, Appl
43	431.5	30.1	300	2	US-10-059-964A-50	Sequence 50, Appl
44	427	29.8	297	2	US-09-314-701-14	Sequence 14, Appl
45	427	29.8	297	2	US-10-314-639-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-09-660-587-46  
; Sequence 46, Application US/09660587  
; Patent No. 6392023  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/660,587  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-09-660-587-46

Query Match	100.0%	Score 1434;	DB 2;	Length 271;
Best Local Similarity	100.0%	Pred. No 2.3e-147;	Mismatches 0;	Indels 0; Gaps 0;
Matches 271;	Conservative 0;			
Qy	1	MNYKRFVVGVTSTFVFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFNFSAEETIPG	60	
Db	1	MNYKRFVVGVTSTFVFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFNFSAEETIPG	60	
Qy	61	ITKKIFALGDKSEINTHSNFTSYDPTASSPAGSGIIGYVNDFRVEFGSYENPEP	120	
Db	61	ITKKIFALGDKSEINTHSNFTSYDPTASSPAGSGIIGYVNDFRVEFGSYENPEP	120	
Qy	121	EROWYPENSOSYKFFALSRLNATNSDNKFI VLENNGVVDKSLNNVVCYDIASGSIPLAPYM	180	
Db	121	EROWYPENSOSYKFFALSRLNATNSDNKFI VLENNGVVDKSLNNVVCYDIASGSIPLAPYM	180	
Qy	181	CAGVGADYIKFLGISLSPKFSYQVKFGVYVNVNNTMLFGGGYHKVVGDRHERVEIAVHP	240	
Db	181	CAGVGADYIKFLGISLSPKFSYQVKFGVYVNVNNTMLFGGGYHKVVGDRHERVEIAVHP	240	
Qy	241	TALSDVPRTTTSASATLNTDYFGWEIGRPFAL	271	
Db	241	TALSDVPRTTTSASATLNTDYFGWEIGRPFAL	271	

RESULT 2

US-09-811-007A-46  
; Sequence 46, Application US/09811007A  
; Patent No. 6660269  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/811,007A  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-09-811-007A-46

Query Match 100.0%; Score 1434; DB 2; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2,3e-147;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEETIPG 60  
DB 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEETIPG 60  
QY 61 ITKKIFALGDKSEINTHSNFTSYDPTYASSFAGSGIIGYVNDPRVFEFGSYENFEP 120  
DB 61 ITKKIFALGDKSEINTHSNFTSYDPTYASSFAGSGIIGYVNDPRVFEFGSYENFEP 120  
QY 121 ERQWYSPNSQSYKFFALSRNATNSDNKFI VLENNGVVVDKSLNVNVCYDIASGSIPLAPYM 180  
DB 121 ERQWYSPNSQSYKFFALSRNATNSDNKFI VLENNGVVVDKSLNVNVCYDIASGSIPLAPYM 180  
QY 181 CAGVGADYIKFLGISLPKFSYQVKGFGVNYPLNVNMTLFGGGYHKVVGDRHVERVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPKFSYQVKGFGVNYPLNVNMTLFGGGYHKVVGDRHVERVEIAYHP 240  
QY 241 TALSDVPRTTSSASATLNTDYFGWEIGRFPAL 271  
DB 241 TALSDVPRTTSSASATLNTDYFGWEIGRFPAL 271

RESULT 3  
US-09-660-587-42  
; Sequence 42, Application US/09660587  
; Patent No. 6392023  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/660,587  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 42  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein  
US-09-660-587-42

Query Match 35.5%; Score 509.5; DB 2; Length 280;

Best Local Similarity 40.5%; Pred. No. 7.5e-47;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
QY 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLSILPYQSFPADPVGSRNDNKEGYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIFALGDKSEINTHSNFT---SYDPTYASSFAGSGIIGYVNDPR 108  
DB 61 PINGTNSLTQKVFGLKKD-GDITKKDDTRVAPGIDFQNNLIS---GFSGSIGYSMDGPR 116  
QY 109 VEFEGSYENFEPER--QWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVVDKSLNVNVC 166  
DB 117 IELEAAYQQFNPKNNTDNDTNGEYKHFALSRKDAMEDQQVVLKNDGITFMSLMVNTC 176  
QY 167 YDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKGFGVNYPLNVNMTLFGGGYHK 225  
DB 177 YDITAEGVSVFPYACAGIGADLITIFKDLNL-KFAYQKGIGISYDITPEVSFAFIGGYHG 235  
QY 226 VVGDRHVERVEIAYHPTALSDVPRTTSSASATLNTDYFGWEIGRFP 269  
DB 236 VIGNKFEKIPV-ITPVLNDAPQTTASVTLVDVGYPFGGEIGMRP 278

RESULT 4  
US-09-314-701-48  
; Sequence 48, Application US/09314701  
; Patent No. 6544517  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko  
; APPLICANT: Ohasi, No. 6544517io  
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
; TITLE OF INVENTION: Chaffeensis  
; FILE REFERENCE: 22727/04021  
; CURRENT APPLICATION NUMBER: US/09/314,701  
; CURRENT FILING DATE: 1999-05-19  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
US-09-314-701-48

Query Match 35.5%; Score 509.5; DB 2; Length 280;  
Best Local Similarity 40.5%; Pred. No. 7.5e-47;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

QY 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLSILPYQSFPADPVGSRNDNKEGYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIFALGDKSEINTHSNFT---SYDPTYASSFAGSGIIGYVNDPR 108  
DB 61 PINGTNSLTQKVFGLKKD-GDITKKDDTRVAPGIDFQNNLIS---GFSGSIGYSMDGPR 116  
QY 109 VEFEGSYENFEPER--QWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVVDKSLNVNVC 166  
DB 117 IELEAAYQQFNPKNNTDNDTNGEYKHFALSRKDAMEDQQVVLKNDGITFMSLMVNTC 176  
QY 167 YDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKGFGVNYPLNVNMTLFGGGYHK 225  
DB 177 YDITAEGVSVFPYACAGIGADLITIFKDLNL-KFAYQKGIGISYDITPEVSFAFIGGYHG 235  
QY 226 VVGDRHVERVEIAYHPTALSDVPRTTSSASATLNTDYFGWEIGRFP 269  
DB 236 VIGNKFEKIPV-ITPVLNDAPQTTASVTLVDVGYPFGGEIGMRP 278

RESULT 5  
US-09-811-007A-42  
; Sequence 42, Application US/09811007A  
; Patent No. 6660269



/ APPLICANT: Walker, David H.  
/ APPLICANT: McBride, Jere W.  
/ APPLICANT: Yu, Xue-Jie  
/ TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
/ TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
/ FILE REFERENCE: D6152CIP2  
/ CURRENT APPLICATION NUMBER: US/09/660,587  
/ PRIOR FILING DATE: 2000-09-12  
/ PRIOR APPLICATION NUMBER: 09/261,358  
/ PRIOR FILING DATE: 1999-03-03  
/ NUMBER OF SEQ ID NOS: 46  
/ SEQ ID NO 10  
/ LENGTH: 283  
/ TYPE: PRT  
/ ORGANISM: Ehrlichia chaffeensis  
/ FEATURE:  
/ OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1b  
US-09-660-587-10

Query Match 35.4%; Score 507; DB 2; Length 283;  
Best Local Similarity 40.3%; Pred. No. 1.4e-46;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

Qy 1 MNYKRFVVGVTLSFVFFFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54  
Db 1 MNYKKIFVSSALISLSILPYQSFADPVTSDTNGINDSREGFYISVKYNPSISHFRKFA 60  
Qy 55 EET-IPG---ITKKIFALGLDKSEINHTSNFTSRSDPTYA---SSPAGSGIIGYVNDP 107  
Db 61 EAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGISYAMDGP 118  
Qy 108 RVFEFGSYENFPERQWYPENSOS-----YKFPALSRNATNSDNKFIVLNNGVVDKSLN 162  
Db 119 RIELEAAVQKFDAXN---PDNDTNSGDYKYFGLSREDAIADKKVVLKNEGITFMSLM 175  
Qy 163 VNVCDYIAGSISPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPNVTMLFGGG 221  
Db 176 VNTCYDITAEGVFPFIPYACAGVGADLINVKDFNL-KFSYQKGIGISYIPITPEVSFIG 234  
Qy 222 YHKVVGDRHVERVEIAYHPTALSDVPRTTSSATLNTDYFGWEIGFRF 269  
Db 235 YYHGVIGNNFNFKIPV-ITPVVLEGAPQTTLSALVTIDTGYFGGEGVGRF 281

RESULT 9  
US-09-261-358A-10  
/ Sequence 10, Application US/09261358A  
/ Patent No. 6403780  
/ GENERAL INFORMATION:  
/ APPLICANT: Walker, David H.  
/ APPLICANT: McBride, Jere W.  
/ APPLICANT: Yu, Xue-Jie  
/ TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
/ TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
/ FILE REFERENCE: D6152CIP  
/ CURRENT APPLICATION NUMBER: US/09/261,358A  
/ CURRENT FILING DATE: 1999-03-03  
/ PRIOR APPLICATION NUMBER: 09/201,458  
/ PRIOR FILING DATE: 1998-11-30  
/ NUMBER OF SEQ ID NOS: 33  
/ SEQ ID NO 10  
/ LENGTH: 283  
/ TYPE: PRT  
/ ORGANISM: Ehrlichia chaffeensis  
/ FEATURE:  
/ OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1b  
US-09-261-358A-10

Query Match 35.4%; Score 507; DB 2; Length 283;  
Best Local Similarity 40.3%; Pred. No. 1.4e-46;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

Qy 1 MNYKRFVVGVTLSFVFFFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54

Db 1 MNYKKIFVSSALISLSILPYQSFADPVTSDTNGINDSREGFYISVKYNPSISHFRKFA 60  
Qy 55 EET-IPG---ITKKIFALGLDKSEINHTSNFTSRSDPTYA---SSPAGSGIIGYVNDP 107  
Db 61 EAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGISYAMDGP 118  
Qy 108 RVFEFGSYENFPERQWYPENSOS-----YKFPALSRNATNSDNKFIVLNNGVVDKSLN 162  
Db 119 RIELEAAVQKFDAXN---PDNDTNSGDYKYFGLSREDAIADKKVVLKNEGITFMSLM 175  
Qy 163 VNVCDYIAGSISPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPNVTMLFGGG 221  
Db 176 VNTCYDITAEGVFPFIPYACAGVGADLINVKDFNL-KFSYQKGIGISYIPITPEVSFIG 234  
Qy 222 YHKVVGDRHVERVEIAYHPTALSDVPRTTSSATLNTDYFGWEIGFRF 269  
Db 235 YYHGVIGNNFNFKIPV-ITPVVLEGAPQTTLSALVTIDTGYFGGEGVGRF 281

RESULT 10  
US-09-201-458-6  
/ Sequence 6, Application US/09201458A  
/ Patent No. 6458942  
/ GENERAL INFORMATION:  
/ APPLICANT: Walker, David H.  
/ APPLICANT: McBride, Jere W.  
/ APPLICANT: Yu, Xue-Jie  
/ TITLE OF INVENTION: 28-kDa Immunoreactive Protein Gene of Ehrlichia  
/ TITLE OF INVENTION: canis and Uses Thereof  
/ FILE REFERENCE: D6152  
/ CURRENT APPLICATION NUMBER: US/09/201,458A  
/ CURRENT FILING DATE: 1998-11-30  
/ NUMBER OF SEQ ID NOS: 21  
/ SEQ ID NO 6  
/ LENGTH: 283  
/ TYPE: PRT  
/ ORGANISM: Ehrlichia chaffeensis  
/ FEATURE:  
/ OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1b  
US-09-201-458-6

Query Match 35.4%; Score 507; DB 2; Length 283;  
Best Local Similarity 40.3%; Pred. No. 1.4e-46;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

Qy 1 MNYKRFVVGVTLSFVFFFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54  
Db 1 MNYKKIFVSSALISLSILPYQSFADPVTSDTNGINDSREGFYISVKYNPSISHFRKFA 60  
Qy 55 EET-IPG---ITKKIFALGLDKSEINHTSNFTSRSDPTYA---SSPAGSGIIGYVNDP 107  
Db 61 EAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGISYAMDGP 118  
Qy 108 RVFEFGSYENFPERQWYPENSOS-----YKFPALSRNATNSDNKFIVLNNGVVDKSLN 162  
Db 119 RIELEAAVQKFDAXN---PDNDTNSGDYKYFGLSREDAIADKKVVLKNEGITFMSLM 175  
Qy 163 VNVCDYIAGSISPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPNVTMLFGGG 221  
Db 176 VNTCYDITAEGVFPFIPYACAGVGADLINVKDFNL-KFSYQKGIGISYIPITPEVSFIG 234  
Qy 222 YHKVVGDRHVERVEIAYHPTALSDVPRTTSSATLNTDYFGWEIGFRF 269  
Db 235 YYHGVIGNNFNFKIPV-ITPVVLEGAPQTTLSALVTIDTGYFGGEGVGRF 281

RESULT 11  
US-09-314-701-4  
/ Sequence 4, Application US/09314701  
/ Patent No. 6544517  
/ GENERAL INFORMATION:  
/ APPLICANT: Rikihisa, Yasuko





```
; CURRENT FILING DATE: 2002-01-28
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 4
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; US-10-059-964A-4

Query Match          35.4%; Score 507; DB 2; Length 283;
Best Local Similarity 40.3%; Pred. No. 1.4e-46;
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

Qy 1 MNYKRFVGVGTLSTFVFFLSDGAF-----SDANFSEGRRLGTYGQYKVGIPNFSNFA 54
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MNYKKIPVSSALISLMSILPYQSFADPVTSDTGINDSREGFYISVKYNPSISHFKESA 60

Qy 55 EET-IPG---ITKKIFALGLDKSEINTHSNFTRSYDPTYA---SSPAGSGLIIGYYVND 107
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 EEAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGSIGYAMDGP 118

Qy 108 RVFEFGSYENFERQWYPENSOS-----YKFFALSRNATNSDNKFTVLENNGVVDKSLN 162
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 RILEAAYQFDKDN---PDNNDTNSGDYKYFGLSKREDIAIDKKYVVLKNEGITFMSLM 175

Qy 163 VNVCYDIAGSISPLAPYMCAGVGADYIK-FLGISLPKFSYQVKVGVNYPINVTMLFGGG 221
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 VNTCYDITASGVFPPIPVACAGVADLINVFKDNL-KFSYQKGIGISYPTITPEVSFIGG 234

Qy 222 YVHKVVDGRHERVEIAVHPTALSDVPRTTTSASATLNTDYFGWEIGRFR 269
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 YVHGVIQNNFNKIPV-ITPVVLEGAPQTTTSALVTIDTGYFGGSGVGRF 281
```

```
RESULT 15
US-09-314-701-30
; Sequence 30, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; FILE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 30
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; US-09-314-701-30

Query Match          34.3%; Score 492; DB 2; Length 285;
Best Local Similarity 38.4%; Pred. No. 6.1e-45;
Matches 111; Conservative 56; Mismatches 96; Indels 26; Gaps 10;

Qy 1 MNYKR--FVGVVTLSTFVFFLSDGAFSD--ANFSEGRRLGTYGQYKVGIPNFSNFAEE 56
    || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MNNRKFFIIGSLASLLFTSEASTGNSGVNSNHTYKPRLYISGQYRPGVSHFSKFSVKE 60

Qy 57 T-----IPGITKKIFALGLDKSEINTHSNFTRSYDPTYASSPAGSGLIIGY-YVNDFRV 109
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 TNYNTQLVGLKKDLSVIG--NSNITYTNFNPYIAEFQDNALISFGAIGLYISENFRI 118

Qy 110 EFEGSVENFERQWYPENS--OSYKFFALSR-----NATNSDN--KFIVLENNGVVDK 160
    || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 EVEASYEEFDVKN---PEGSATDAYRYFALARAMDGTNKSPPDTRKFTVMRNDGLSISS 175
```

```
Qy 161 LNNVVCYDIAGSISPLAPYMCAGVGADYIKFLGISLPKFSYQVKGVNYPINVTMLFGG 220
Db 176 VMINGCYNFTLDDIPVVPYVCAGIGGDFIEFFNDLHVKFRHQKVGISYSISPEVSLFLN 235
Qy 221 GYVHKVVDGRHERVEIAVHPTALSDVPRTTTSASATLNTDYFGWEIGRFR 269
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 GYVHKVYGNRFXNLHV-QHVSDLSADAPKFTSAVATLNVGYFGGEIGVRF 283
```

Search completed: December 28, 2005, 15:43:04  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:02:01 ; Search time 111 Seconds  
(without alignments)

1722.506 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNYKRFVGVGTLSTFVFLS.....ASATLNTDYFGWEIGFRPAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80.\*

1: uniprot\_eprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1434	100.0	271	2	Q9F471_EHRCA	Q9F471 ehrlichia c
2	1424	99.3	275	2	Q9ADV2_EHRCA	Q9ADV2 ehrlichia c
3	1194	83.3	271	2	O69197_EHRCH	O69197 ehrlichia c
4	903	63.0	283	2	O5FGQ6_EHRRG	O5FGQ6 ehrlichia r
5	903	63.0	283	2	Q4LOB8_EHRRU	Q4LOB8 ehrlichia r
6	892	62.2	285	2	Q6VCX5_EHRRU	Q6VCX5 ehrlichia r
7	892	62.2	285	2	Q5HA08_EHRRW	Q5HA08 ehrlichia r
8	856	59.7	271	2	Q4W4W3_EHRRU	Q4W4W3 ehrlichia r
9	525.5	36.6	282	2	O5W5Y1_EHRRU	O5W5Y1 ehrlichia r
10	522.5	36.4	282	2	Q5HA10_EHRRU	Q5HA10 ehrlichia r
11	522.5	36.4	282	2	Q9W4W1_EHRRU	Q9W4W1 ehrlichia r
12	517.5	36.1	282	2	Q9R443_EHRRU	Q9R443 ehrlichia r
13	515	35.9	283	2	Q8G812_EHRCH	Q8G812 ehrlichia c
14	514.5	35.9	282	2	Q4LOC0_EHRRU	Q4LOC0 ehrlichia r
15	514.5	35.9	304	2	O5FGL4_EHRRG	O5FGL4 ehrlichia r
16	514.5	35.9	304	2	O5FCD9_EHRRW	O5FCD9 ehrlichia r
17	514	35.8	283	2	Q8G8D6_EHRCH	Q8G8D6 ehrlichia c
18	509.5	35.5	280	2	Q9F476_EHRRU	Q9F476 ehrlichia c
19	507	35.4	283	2	O8GGU4_EHRRU	O8GGU4 ehrlichia c
20	507	35.4	283	2	O85358_EHRRU	O85358 ehrlichia c
21	504.5	35.2	280	2	Q84HU1_EHRRU	Q84HU1 ehrlichia c
22	494	34.4	285	2	O9L6V5_EHRRU	O9L6V5 ehrlichia c
23	489	34.1	279	2	Q9ACJ2_EHRRU	Q9ACJ2 ehrlichia c
24	482	33.6	279	2	Q9L6Y9_EHRRU	Q9L6Y9 ehrlichia c
25	469	32.7	282	2	Q4W4X0_EHRRU	Q4W4X0 ehrlichia r
26	467	32.6	282	2	O5FGM2_EHRRG	O5FGM2 ehrlichia r
27	467	32.6	282	2	Q4LOC7_EHRRU	Q4LOC7 ehrlichia r
28	464	32.4	282	2	O5HA16_EHRRW	O5HA16 ehrlichia r
29	463	32.3	283	2	Q9L6Y7_EHRRU	Q9L6Y7 ehrlichia c
30	460	32.1	281	2	Q9ADV7_EHRRU	Q9ADV7 ehrlichia c
31	460	32.1	282	2	Q6VCY4_EHRRU	Q6VCY4 ehrlichia r

32	455.5	31.8	298	2	Q9L6V3_EHRCH	Q9L6V3 ehrlichia c
33	447.5	31.2	300	2	Q9L6V2_EHRCH	Q9L6V2 ehrlichia c
34	446	31.1	289	2	Q9ADV9_EHRCA	Q9ADV9 ehrlichia c
35	442.5	30.9	294	2	Q9ADV4_EHRCA	Q9ADV4 ehrlichia c
36	442	30.8	278	2	Q9ADW1_EHRCA	Q9ADW1 ehrlichia c
37	440	30.7	291	2	Q9ACJ0_EHRCH	Q9ACJ0 ehrlichia c
38	437.5	30.5	290	2	O5FGL6_EHRRG	O5FGL6 ehrlichia r
39	437.5	30.5	290	2	O4LOC1_EHRRU	O4LOC1 ehrlichia r
40	437	30.5	293	2	Q9F477_EHRCA	Q9F477 ehrlichia c
41	436	30.4	275	2	Q93DD4_EHRCH	Q93DD4 ehrlichia c
42	435.5	30.4	288	2	Q9ZGJ2_EHRCA	Q9ZGJ2 ehrlichia c
43	428	29.8	283	2	O4LOC6_EHRRU	O4LOC6 ehrlichia r
44	428	29.8	283	2	O6VCY3_EHRRU	O6VCY3 ehrlichia r
45	428	29.8	283	2	Q5HA15_EHRRW	Q5HA15 ehrlichia r

ALIGNMENTS

RESULT 1  
Q9F471\_EHRCA  
ID Q9F471\_EHRCA PRELIMINARY; PRT; 271 AA.  
AC Q9F471;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE P28-9.  
GN Name=p28-9;  
OS Ehrlichia canis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Anaplasmataceae; Ehrlichia.  
OX NCBI\_TaxID=944;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=JAKE;  
RX MEDLINE=99242757; PubMed=10225842;  
RA McBride J.W., Yu X.J., Walker D.H.;  
RT "Molecular cloning of the gene for a conserved major immunoreactive  
RT 28-kilodalton protein of Ehrlichia canis: a potential serodiagnostic  
RT antigen."  
RL Clin. Diagn. Immunol. 6:392-399(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=JAKE;  
RX MEDLINE=20432107; PubMed=10974556; DOI=10.1016/S0378-1119(00)00256-0;  
RA McBride J.W., Yu X.J., Walker D.H.;  
RT "A conserved, transcriptionally active p28 multigene locus of  
RT Ehrlichia canis."  
RL Gene 254:245-252(2000).  
DR EMBL; AF082744; AAG14363.1; -; Genomic DNA.  
DR InterPro; IPR002566; Surface\_Ag\_msp4.  
DR Pfam; PF01617; Surface\_Ag\_2; 1.  
SQ SEQUENCE 271 AA; 30374 MW; 8C8F22BB1D05B3EE CRC64;

Query Match 100.0%; Score 1434; DB 2; Length 271;  
Best Local Similarity 100.0%; Pred. No. 1.2e-109;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNYKRFVGVGTLSTFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60  
DB 1 MNYKRFVGVGTLSTFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60  
QY 61 ITTKIFALGLDKSEINHSNTRSDPTDYASSFAGSGIIGYVNDFRVFEFGSVENFPP 120  
DB 61 ITTKIFALGLDKSEINHSNTRSDPTDYASSFAGSGIIGYVNDFRVFEFGSVENFPP 120  
QY 121 ERQWTPENSQSYKFFALSRNATNSDNKFTVLNNGVVDKSLNNVCYDIASGSIPLAPYM 180  
DB 121 ERQWTPENSQSYKFFALSRNATNSDNKFTVLNNGVVDKSLNNVCYDIASGSIPLAPYM 180  
QY 181 CAGVGADYIKFGLIGISLPKFSQVKGAVNPLNVTMLFGGYYHKVGVGRHVERVEYHP 240  
DB 181 CAGVGADYIKFGLIGISLPKFSQVKGAVNPLNVTMLFGGYYHKVGVGRHVERVEYHP 240

```

QY 241 TALSVPRTTSASATLNTDYFGWEIGFRFAL 271
DB 241 TALSVPRTTSASATLNTDYFGWEIGFRFAL 271

RESULT 2
Q9ADV2_EHRCA PRELIMINARY; PRT; 275 AA.
AC Q9ADV2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Major outer membrane protein P30-20.
GN Name=p30-20;
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=944;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Oklahoma;
RC MEDLINE=98371112; PubMed=9705412;
RX DOI=10.1128/IAI.69.4.2083-2091.2001;
RA Ohashi N., Unver A., Zhi N., Rikihisa Y.;
RT "Cloning and characterization of multigenes encoding the
RT immunodominant 30-kilodalton major outer membrane proteins of
RT Ehrlichia canis and application of the recombinant protein for
RT serodiagnosis.";
RL J. Clin. Microbiol. 36:2671-2680 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oklahoma;
RX MEDLINE=21153566; PubMed=11254561;
RX DOI=10.1128/IAI.69.4.2083-2091.2001;
RA Ohashi N., Rikihisa Y., Unver A.;
RT "Analysis of transcriptionally active gene clusters of major outer
RT membrane protein multigene family in Ehrlichia canis and E.
RT chaffeensis.";
RL Infect. Immun. 69:2083-2091 (2001).
DR EMBL; AF078553; AAK28700.1; -; Genomic_DNA.
DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 275 AA; 30843 MW; 3D8FD88FE507CE51 CRC64;

Query Match 99.3%; Score 1424; DB 2; Length 275;
Best Local Similarity 99.3%; Pred. No. 7.8e-109;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYKRFVGVVLTSTFVFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
DB 5 MNYKRFVGVVLTSTFVFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 64
QY 61 ITKKIFALGLDKSEINTHSNFTSRSDPTYASSPAGFSGIIGYVNDPRVFEFGSYENFEP 120
DB 65 ITKKIFALGLDKSEINTHSNFTSRSDPTYASSPAGFSGIIGYVNDPRVFEFGSYENFEP 124
QY 121 ERQWYPENSQSKFFALSRLNATSDNKFIVLENNGVVDKSLNNVVCYDIASGSIPLAYYM 180
DB 125 ERQWYPENSQSKFFALSRLNATSDNKFIVLENNGVVDKSLNNVVCYDIASGSIPLAYYM 184
QY 181 CAGVGADYIKFLGISLPKFSYQVKGNYPLNVNTMLFGGGYHKVGDHRRVERVEIAYHP 240
DB 185 CAGVGADYIKFLGISLPKFSYQVKGNYPLNVNTMLFGGGYHKVGDHRRVERVEIAYHP 244
QY 241 TALSVPRTTSASATLNTDYFGWEIGFRFAL 271
DB 245 TALSVPRTTSASATLNTDYFGWEIGFRFAL 275

RESULT 3
O69197_EHRCH PRELIMINARY; PRT; 271 AA.
AC O69197;

QY 241 TALSVPRTTSASATLNTDYFGWEIGFRFAL 271
DB 241 TALSVPRTTSASATLNTDYFGWEIGFRFAL 271

RESULT 4
Q5FGQ6_EHRRG PRELIMINARY; PRT; 283 AA.
AC Q5FGQ6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)

```

RESULT 7  
Q5HA08 EHRRW  
ID Q5HA08 EHRRW PRELIMINARY;  
PRT; 285 AA.

```
AC Q5HA08; Q5FCD7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative outer membrane protein Mapi-1 (Mapi-related protein).
GN Name=mapi-1; OrderedLocusNames=ERWE_CDS_09260, Brum8750;
OS Ehrlichia ruminantium (strain Welgevonden).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=254945;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15637156; DOI=10.1073/pnas.0406633102;
RA Collins N.E., Liebenberg J., de Villiers E.P., Brayton K.A., Louw E.,
RA Pretorius A., Faber F.E., van Heerden H., Josenmans A., van Kleef M.,
RA Steyn H.C., van Strijp M.F., Zweygardh E., Jongejan F., Maillard J.C.,
RA Berthier D., Botha M., Joubert F., Corton C.H., Thomson N.R.,
RA Allsopp M.T., Allsopp B.A.;
RT "The genome of the heartwater agent Ehrlichia ruminantium contains
RT multiple tandem repeats of actively variable copy number.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:838-843(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Frutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y., Viari A.,
RA Chantal I., Morgat A., Colisac E., Vachery N., Demaille J., Viari A.,
RA Martinez D.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR767821; CAH58610.1; -; Genomic DNA.
DR EMBL; CR925678; CA127420.1; -; Genomic DNA.
DR InterPro; IPR002566; Surface_Ag_map4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ Complete proteome.
SQ SEQUENCE 285 AA; 32108 MW; 05E4C4E59A6F9118 CRC64;

Query Match 62.2%; Score 892; DB 2; Length 285;
Best Local Similarity 59.1%; Pred. No. 4.5e-65;
Matches 166; Conservative 45; Mismatches 60; Indels 10; Gaps 4;

Qy 1 MNYKRFVGVVTLSTVFP-FLSDGAFSDANFSEGRRLGYIGSQYKVGIPNFSNFAEETIP 59
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5 MNYKEFVLGVLSALLFSLPERAISDMVDENSRFRFAGVQYRTGIPNFDNFSASETIP 64

Qy 60 GITKKIFALGLD--KSEINTHSNFTSRVDPTYASSFAGSGIIGYVNDFRVEPEGSVEN 117
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 GLTKGVYGLDLDSKSDITKRAINFRLYNPTYSTSTGIGMGFYGFNIRMEFETSYSS 124

Qy 118 FEPEQWYPENSQVKPFALSRN---ATNSP----NKFIVLNNGVVDKSLNVCYDIA 170
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 FGIERQWYPEGSQSHKFCVSRQDNAPNTDSSNNNDFFVLNNGVKIRTLHVNFCYDVA 184

Qy 171 SGSIPLAPYMCAGVADYIKELGISLPKFSYQVKFGVNYPLNVTMLFGGYYHKVVGDR 230
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 HGNIPLAYPYCAGIGDVVKIGVSLPKFSYQLKFGVNYPLSIRTMLFGGYYHKVWGSK 244

Qy 231 HERVEIAIHPALSDVPRTTSASATLNTDYFGWEIGFRFAL 271
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 YDRVKVYHPVQLNTVPKMTFVSANLIDIDYFCGCEVGIRFFL 285

RESULT 8
Q4W4W3_EHRRU
ID Q4W4W3_EHRRU PRELIMINARY; PRT; 271 AA.
AC Q4W4W3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Mapi-1 (Fragment).
GN Name=mapi-1;
OS Ehrlichia ruminantium (Cowdria ruminantium).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=779;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RX MEDLINE=22035368; PubMed=12039046; DOI=10.1016/S0378-1119(02)00408-0;
RA Bekker C.P., Bell-Sakyi L., Paxton E.A., Martinez D., Bensaid A.,
RA Jongejan F.;
RT "Transcriptional analysis of the major antigenic protein 1 multigene
RT family of Cowdria ruminantium.";
RL Gene 285:193-201(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RX PubMed=1595193; DOI=10.1128/JB.187.14.4782-4791.2005;
RA Bekker C.P., Postigo M., Taoufik A., Bell-Sakyi L., Ferraz C.,
RA Martinez D., Jongejan F.;
RT "Transcription Analysis of the Major Antigenic Protein 1 Multigene
RT Family of Three In Vitro-Cultured Ehrlichia ruminantium Isolates.";
RL J. Bacteriol. 187:4782-4791(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RA Bekker C.P.J., Taoufik A., Jongejan F.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP319940; AAY59059.1; -; Genomic DNA.
FT NON_TER 1
SQ SEQUENCE 271 AA; 30060 MW; F6AB82905DAA7690 CRC64;

Query Match 59.7%; Score 856; DB 2; Length 271;
Best Local Similarity 59.0%; Pred. No. 3.9e-62;
Matches 160; Conservative 42; Mismatches 61; Indels 8; Gaps 3;

Qy 9 GVTLSSTVFP-FLSDGAFSDANFSEGRRLGYIGSQYKVGIPNFSNFAEETIPGITKKIFA 67
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 GVTLSALLFSLPERAISDMVDENSRFRFAGVQYRTGIPNFDNFSASETIPGLTKGVY 60

Qy 68 LGLD--KSEINTHSNFTSRVDPTYASSFAGSGIIGYVNDFRVEPEGSVENPEPEQWY 125
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LDLDLSKSGITKRAINFRLYNPTYSTSTGIGMGFYGFNIRMEFETSYSNFGVERQWY 120

Qy 126 PENSQSKYKPFALSRNAT----NSDNKFIVLNNGVVDKSLNVCYDIASGSIPLAPYM 180
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 PEGSQSHKFCVSRNATPTDSSNNNDFFVLNNGVKIRTLNVCYDVAHGNIPLAPYV 180

Qy 181 CAGVGADYIKELGISLPKFSYQVKFGVNYPLNVTMLFGGYYHKVVGDRHERVEIAIHP 240
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 CAGIGGDYVYKFIGVSLPKFSYQLKFGVNYPLSIRTMLFGGYYHKVYKVDYRVKSSL 240

Qy 241 TALSDVPRTTSASATLNTDYFGWEIGFRFAL 271
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 VQLSTVPKMTFVSANLIDIDYFCGCEVGIRFFL 271

RESULT 9
Q5W5Y1_EHRRU
ID Q5W5Y1_EHRRU PRELIMINARY; PRT; 282 AA.
AC Q5W5Y1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Mapi-1.
GN Name=mapi-1;
OS Ehrlichia ruminantium (Cowdria ruminantium).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=779;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RX MEDLINE=22035368; PubMed=12039046; DOI=10.1016/S0378-1119(02)00408-0;
RA Bekker C.P., Bell-Sakyi L., Paxton E.A., Martinez D., Bensaid A.,
RA Jongejan F.;
RT "Transcriptional analysis of the major antigenic protein 1 multigene
RT family of Cowdria ruminantium.";
RL Gene 285:193-201(2002).
```

```

RN RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Senegal;
RX PubMed=15995193; DOI=10.1128/JB.187.14.4782-4791.2005;
RA Bekker C.P., Postigo M., Taoufik A., Bell-Saki L., Ferraz C.,
RA Martinez D., Jongejan F.;
RT "Transcription Analysis of the Major Antigenic Protein 1 Multigene
RT Family of Three in Vitro-Cultured Ehrlichia ruminantium Isolates.";
RL J. Bacteriol. 187:4782-4791 (2005).
RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RA Bekker C.P., Jongejan A., Jongejan F.;
RA Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF319940; AAV54088.2; -; Genomic DNA.
SQ SEQUENCE 282 AA; 31097 MW; 3D619510E7393EBB CRC64;

Query Match 36.6%; Score 525.5; DB 2; Length 282;
Best Local Similarity 41.8%; Pred. No. 6.1e-35;
Matches 120; Conservative 47; Mismatches 95; Indels 25; Gaps 10;

QY 1 MNYKFFVGVVTLSTFVFFLSDGAFSDANFS-----GRRGLYIGSQYKVGIPNFSNFA 54
DB 1 MNYKILVRSALISLMSFLPYQSPAEPVSSNNIGNEXAKEGFIYSAKYNPSIPHRKESA 60

QY 55 EET----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASS---FAGSGIIGYYVND 107
DB 61 EETPVYKDSPTKKVFGKKGGS-ITKYSDFTRT-DISFEGQNNFISGSGSIGYIMDGP 118

QY 108 RVEFGSYENPEPERQWYPEN-----SQSYKFPALSRNATNSDNKFTVLENNGVVDKSLNV 163
DB 119 RVEIEAAYQKFNPKN---PANETDSDYYKHGYSRAETMTDKKYVVLTVNNGVTSSLMF 175

QY 164 NVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGY 222
DB 176 NACYDITAEVGFPIYACAGIADLISIFDDINL-KFAYQKIGISYPTTPEISAFIGY 234

QY 223 YHKVGVDRHERVEIAYHPTALSDVPRTTSSATLNTDYFGWEIGFRF 269
DB 235 YHGVIQNKYKVPVKL-PVTLTDAPOSTASVTLTDAGYFGGELGVRF 280

RESULT 10
QSHA10 EHRRW PRELIMINARY; PRT; 282 AA.
AC QSHA10;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Putative outer membrane protein MAPI-1.
GN Name=map1-1; OrderedLocusNames=Erum8730;
OS Ehrlichia ruminantium (strain Welgevonden).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=254945;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=15637156; DOI=10.1073/pnas.0406631102;
RX Collins N.E., Liebenberg J., de Villiers E.P., Brayton K.A., Louw E.,
RA Pretorius A., Faber P.E., van Heerden H., Josemans A., van Kleef M.,
RA Steyn H.C., van Strijp M.P., Zweygath E., Jongejan F., Maillard J.C.,
RA Berthier D., Botha M., Joubert F., Corton C.H., Thomson N.R.,
RA Allsopp M.T., Allsopp B.A.;
RT "The genome of the heartwater agent Ehrlichia ruminantium contains
RT multiple tandem repeats of actively variable copy number.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:838-843 (2005).
DR EMBL; CR767821; CAH58608.1; -; Genomic DNA.
DR InterPro; IPR011539; RHD.
DR InterPro; IPR002566; Surface Ag msp4.
DR InterPro; IPR011991; Wing hlx DNA_bd.
DR Pfam; PF01617; Surface_Ag_2; 1.
KW Complete proteome.
SQ SEQUENCE 282 AA; 31100 MW; 1CB2DA256B2CA990 CRC64;

Query Match 36.4%; Score 522.5; DB 2; Length 282;
Best Local Similarity 41.8%; Pred. No. 1.1e-34;
Matches 120; Conservative 47; Mismatches 95; Indels 25; Gaps 10;

QY 1 MNYKFFVGVVTLSTFVFFLSDGAFSDANFS-----EGRRGLYIGSQYKVGIPNFSNFA 54
DB 1 MNYKILVRSALISLMSFLPYQSPAEPVSSNNIGNENAKEGFIYSAKYNPSIPHRKESA 60

QY 55 EET----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASS---FAGSGIIGYYVND 107
DB 61 EETPVYKDSPTKKVFGKKGGS-ITKYSDFTRT-DISFEGQNNFISGSGSIGYIMDGP 118

QY 108 RVEFGSYENPEPERQWYPEN-----SQSYKFPALSRNATNSDNKFTVLENNGVVDKSLNV 163
DB 119 RVEIEAAYQKFNPKN---PANETDSDYYKHGYSRAETMTDKKYVVLTVNNGVTSSLMF 175

QY 164 NVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGY 222
DB 176 NACYDITAEVGFPIYACAGIADLISIFDDINL-KFAYQKIGISYPTTPEISAFIGY 234

QY 223 YHKVGVDRHERVEIAYHPTALSDVPRTTSSATLNTDYFGWEIGFRF 269
DB 235 YHGVIQNKYKVPVKL-PVTLTDAPOSTASVTLTDAGYFGGELGVRF 280

RESULT 11
QSHA10 EHRRW PRELIMINARY; PRT; 282 AA.
AC QSHA10;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Major antigenic protein 1 like protein (MAP1-1).
OS Ehrlichia ruminantium (Cowdria ruminantium).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=779;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=99216274; PubMed=10198207; DOI=10.1006/bbrc.1999.0459;
RX Sulzsa C.R., Mahan S.M., Barbet A.P.;
RT "The map1 gene of Cowdria ruminantium is a member of a multigene
RT family containing both conserved and variable genes.";
RL Biochem. Biophys. Res. Commun. 257:300-305 (1999).
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Welgevonden;
RC van Heerden H., Collins N.E., Brayton K.A., Rademeyer C.,
RA Allsopp B.A.;
RT "Characterization of a major outer membrane protein multigene family
RT in Ehrlichia ruminantium.";
RL Gene 330:159-168 (2004).
DR EMBL; AF125276; AAD26347.1; -; Genomic DNA.
DR EMBL; AF125275; AAD26345.1; -; Genomic DNA.
DR EMBL; AY343331; AARI0944.1; -; Genomic DNA.
DR InterPro; IPR002566; Surface Ag msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 282 AA; 31100 MW; 1CB2DA256B2CA990 CRC64;

Query Match 36.4%; Score 522.5; DB 2; Length 282;
Best Local Similarity 41.8%; Pred. No. 1.1e-34;
Matches 120; Conservative 47; Mismatches 95; Indels 25; Gaps 10;

QY 1 MNYKFFVGVVTLSTFVFFLSDGAFSDANFS-----EGRRGLYIGSQYKVGIPNFSNFA 54
DB 1 MNYKILVRSALISLMSFLPYQSPAEPVSSNNIGNENAKEGFIYSAKYNPSIPHRKESA 60

QY 55 EET----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASS---FAGSGIIGYYVND 107
DB 61 EETPVYKDSPTKKVFGKKGGS-ITKYSDFTRT-DISFEGQNNFISGSGSIGYIMDGP 118

QY 108 RVEFGSYENPEPERQWYPEN-----SQSYKFPALSRNATNSDNKFTVLENNGVVDKSLNV 163
DB 119 RVEIEAAYQKFNPKN---PANETDSDYYKHGYSRAETMTDKKYVVLTVNNGVTSSLMF 175

QY 164 NVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGY 222
DB 176 NACYDITAEVGFPIYACAGIADLISIFDDINL-KFAYQKIGISYPTTPEISAFIGY 234

QY 223 YHKVGVDRHERVEIAYHPTALSDVPRTTSSATLNTDYFGWEIGFRF 269
DB 235 YHGVIQNKYKVPVKL-PVTLTDAPOSTASVTLTDAGYFGGELGVRF 280
```





```
Db      61 EETPVYKDSPTKVKFGLKKEGS-ITKYSDFTRT-DISFEGQNNFISGFSIGYIMDGP 118
Qy      108 RVEFEFGSYNFERQWYPEN-----SQSYKFPALSRNATNSDNKFIVLENNGVVDKSLNV 163
Db      119 RVEIEAAYQKFNPKN---PANETDTSYKHYGLSRAEAMADKKYVVLTNNGVTFSSLMF 175
Qy      164 NVCYDIASGISPLAPYMCAGVADYK-FLGISLPKFSQVKGVYVPLNVNNTMLFGGGY 222
Db      176 NACYDITABGVPIFYACAGIGADLISIPDDINL-KFAYQKGIGISYIPITPEISAFIGGY 234
Qy      223 YHKVVGDRHERVEIAHYHPTALSVPRTTSASATLNTDYFGWEIGPRF 269
Db      235 YHGVIGNKYNKVPVKL-PVTLIDAPQSTSASVTLDAGYFGGELGVRF 280
```

RESULT 15

```
OSFGL4_EHRRG PRELIMINARY; PRT; 304 AA.
AC OSFGL4;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Map1-related protein.
GN OrderedLocusNames=ERGA_CDS_09150;
OS Ehrlichia ruminantium (strain Gardel).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=302409;
RN [1]
RU NUCLEOTIDE SEQUENCE.
RC STRAIN=Gardel;
RA Frutos R., Ferraz C., Bensaïd A., Eychenie S., Kandassami Y.,
RA Chantal I., Morgat A., Coissac E., Vachieri N., Demaille J., Viari A.,
RA Martinez D.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR925677; CAI28367.1; -; Genomic DNA.
DR InterPro; IPR002566; Surface_Ag_mep4.
DR Pfam; PF01617; Surface_Ag_2; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33925 MW; 009338D2B65AE800 CRC64;
```

```
Query Match 35.9%; Score 514.5; DB 2; Length 304;
Best Local Similarity 41.5%; Pred. No. 5.3e-34;
Matches 119; Conservative 46; Mismatches 97; Indels 25; Gaps 10;

Qy      1 MNYKRFVVGVTLSFVFLLSDGAFSDANFS-----EGRRGLYIGSQYKVGIPNFSNRSA 54
Db      23 MNYKILVRSALISLSMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFKKFSA 82

Qy      55 EET- ---IPGITKKIPALGDKSEINTHSNFRSYDPTYASS---PAGEFGIIGYYVNDP 107
Db      83 EETPVYKDSPTKVKFGLKKEGS-ITKYSDFTRT-DISFEGQNNFISGFSIGYIMDGP 140

Qy      108 RVEFEFGSYNFERQWYPEN-----SQSYKFPALSRNATNSDNKFIVLENNGVVDKSLNV 163
Db      141 RVEIEAAYQKFNPKN---PANETDTSYKHYGLSRAEAMADKKYVVLTNNGVTFSSLMF 197

Qy      164 NVCYDIASGISPLAPYMCAGVADYK-FLGISLPKFSQVKGVYVPLNVNNTMLFGGGY 222
Db      198 NACYDITABGVPIFYACAGIGADLISIPDDINL-KFAYQKGIGISYIPITPEISAFIGGY 256

Qy      223 YHKVVGDRHERVEIAHYHPTALSVPRTTSASATLNTDYFGWEIGPRF 269
Db      257 YHGVIGNKYNKVPVKL-PVTLIDAPQSTSASVTLDAGYFGGELGVRF 302
```

Search completed: December 28, 2005, 15:38:00  
Job time : 114 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:05:48 ; Search time 25 Seconds  
(without alignments)  
1042.990 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNYKRFVGVVTLSTVFVFLS.....ASATLNTDYFGWIGFRPAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: piri.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	419.5	29.3	280	JE0217	28k surface antige
2	397.5	27.7	278	JE0216	28k surface antige
3	383.5	26.7	284	I40882	major antigenic pr
4	379	26.4	287	JE0220	28k surface antige
5	378.5	26.4	286	JE0219	28k surface antige
6	377	26.3	276	JE0218	major surface prot
7	278	19.4	282	I39648	28k surface antige
8	146.5	10.2	133	JE0221	iron(III) dicitrat
9	102	7.1	787	G64620	heat resistant aggr
10	100	7.0	284	AG3556	iron (iii) dicitra
11	99	6.9	792	F71894	hypothetical prote
12	97.5	6.8	614	E70120	probable TonB syst
13	97.5	6.8	713	F82506	conserved hypothet
14	97.5	6.8	991	B71315	toxin-like outer m
15	96.5	6.7	1943	B64592	cell surface glyco
16	96	6.7	730	A48832	vacuolating cytoto
17	95	6.6	1287	B53739	hypothetical prote
18	93.5	6.5	3194	D71917	hypothetical prote
19	93	6.5	624	E64483	vacuolating cytot
20	93	6.5	1291	S44983	hypothetical prote
21	92.5	6.5	637	F90257	hypothetical prote
22	92.5	6.5	726	T31887	toxin-like outer m
23	92.5	6.5	2399	H71879	probable outer mem
24	91.5	6.4	239	AH0541	hypothetical prote
25	91.5	6.4	460	T19515	hypothetical prote
26	91.5	6.4	1118	S75309	NADH2 dehydrogenas
27	91	6.3	736	T12716	flagellin chain B
28	90.5	6.3	336	B32808	opacity protein op
29	90	6.3	237	S36343	

30 90 6.3 710 2 C98235 probable hydroxama  
31 90 6.3 716 2 A13050 hydroxamate-type f  
32 90 6.3 1101 2 S58108 hypothetical prote  
33 90 6.3 1769 2 S53378 probable membrane  
34 89 6.2 298 2 B71685 hypothetical prote  
35 89 6.2 482 1 A34671 triacylglycerol li  
36 89 6.2 493 2 G90604 hypothetical prote  
37 88.5 6.2 270 2 A71907 outer membrane pro  
38 88 6.1 797 2 D71621 hypothetical prote  
39 87.5 6.1 896 2 A51514 conserved membrane  
40 87.5 6.1 956 2 T40953 hypothetical prote  
41 87.5 6.1 1532 2 A26039 IGA-specific metal  
42 87.5 6.1 1561 2 S61314 IGA-specific metal  
43 87 6.1 374 2 I40765 hypothetical prote  
44 87 6.1 783 2 I50116 N-cadherin precurs  
45 87 6.1 937 2 S78561 CS3 piliin synthesi

ALIGNMENTS

RESULT 1

JE0217  
28k surface antigen 4 - Ehrlichia chaffensis  
N:Alternate names: MAPI  
C:Species: Ehrlichia chaffensis  
C>Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0217  
R:Reddy, G.R.; Sulsana, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746  
A:Accession: JE0217  
A:Molecule type: DNA  
A:Residues: 1-280 <RED>  
A:Cross-references: UNIPROT:O52107; UNIPARC:UPI000003478E; GB:AF062761

Query Match 29.3%; Score 419.5; DB 2; Length 280;  
Best Local Similarity 34.1%; Pred. No. 1.2e-27;  
Matches 98; Conservative 59; Mismatches 103; Indels 27; Gaps 7;

QY 1 MNYKRFVGVVTLSTVFVFLSDGAFSDANFSEGRG-LVIGSOYKVGIPNFSABETIP 59  
DB 1 MNCKRFFTTTLLVLSMSFLPGISFSDAVQNDVNGNFVIGKVPVSFVSHFGVSAKQE-R 59  
QY 60 GITKKIFAL--GLDKSEINTHSNFRSYDPTVA-----SSFAGSGIIGYVYNDPRVEPE 112  
DB 60 NTTIGVFGKQDWGDSGISKSPENTVNPVYSPKYNPPFLGFAGAVCYLMNGPRIELE 119  
QY 113 GSYENFEPERQWYSPNSQYKFFALSRNA-----TNSDNKFIVLNNGVVDKSLNVNVCYD 168  
DB 120 MSYETFDVKNQGNKYKNDHKYATLTHNSGGKLSNAGDKFVFLKNEGLLDI SILMLNACYD 179  
QY 169 IASGSIPLAPYMCAGVADYIKFLGSLPKFSYQVGVVYPLNVNTMLFGGYVHKVVG 228  
DB 180 VISEGIPSPYICAGVGTDLISMFAINPKISYQKGLSYISPEASVFGVGHFKVIG 239  
QY 229 DRHREIYATHTALSDVPRITSAS-----ATLNTDYFGWIGFRF 269  
DB 240 NEFRDIP-----AMIPSTSLTGNHFTIVTLVSVCHFGVELGGRF 278

RESULT 2

JE0216  
28k surface antigen 3 - Ehrlichia chaffensis  
N:Alternate names: MAPI  
C:Species: Ehrlichia chaffensis  
C>Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0216  
R:Reddy, G.R.; Sulsana, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746

A:Accession: JE0216  
A:Molecule type: DNA  
A:Residues: 1-278 <RED>  
A:Cross-references: UNIPROT:052106; UNIPARC:UPI000003478D; GB:AF062761

Query Match 27.7%; Score 397.5; DB 2; Length 278;  
Best Local Similarity 34.7%; Pred. No. 8.1e-26;  
Matches 99; Conservative 55; Mismatches 106; Indels 25; Gaps 9;

Qy 1 MNYKRFVGVVTLSTFVFLSDGAFSDANFSEGRRLG-LYIGSQYKVGIPNFSNESA-BETI 58  
Db 1 MNCKKFFITVALVLSLFLGSLSDPVGQDINSFNFTVSGKMPFSASHGFMFSKAKEN 60

Qy 59 PGITKKIFAL-----GLDKSEINTHSNFTSYDPTYASS-FAGFSGIIGYVNDFRVEFE 112  
Db 61 P--TVALYGLQDMEGISSSHNDHFNKGVSKYENNPFLGPAAGALGYSMGGPRVEFE 118

Qy 113 GSYENFEPERQWYPENSQYKFFALSRNATN---SDNKFIVLENNGVVDKSLNVNVCYDI 169  
Db 119 VSYETFDVKNQGNKYKNDHRYCALGQODNSGIPKTSKYVLKSEGLLDISFMLNACYDI 178

Qy 170 ASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVNMTLFGGYYHKVGVGD 229  
Db 179 INESIPLSFYCAGVGTDLISMFENATPKISYQKGLGUSYINPEASVFIGGHFKVIGN 238

Qy 230 RHERVEIAYHPTALSDVPRITTSAS-----ATLNTDYFGWEIGGRF 269  
Db 239 EPRDI-----PTLKAFV--TSSATPDLAIVLSVCHFGIELGGRF 276

RESULT 3  
140882  
major antigenic protein - heartwater rickettsia  
C:Species: Cowdria ruminantium (heartwater rickettsia)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: I40882; S42827  
R:van Vliet, A.H.; Jongejan, F.; van Kleef, M.; van der Zeijst, B.A.  
Infect. Immun. 62, 1451-1456, 1994  
A:Title: Molecular cloning, sequence analysis, and expression of the i  
A:Reference number: I40882; MUID:94178956; PMID:8132352  
A:Accession: I40882  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-284 <RES>  
A:Cross-references: UNIPROT:Q46327; UNIPARC:UPI000009A44; EMBL:X74250; NID:g454266; PID  
C:Genetics:  
A:Gene: map1

Query Match 26.7%; Score 383.5; DB 2; Length 284;  
Best Local Similarity 32.9%; Pred. No. 1.2e-24;  
Matches 93; Conservative 55; Mismatches 120; Indels 15; Gaps 7;

Qy 1 MNYKRFVGVVTLSTFVFLSDGAFSDANFSEGR--RGLYIGSQYKVGIPNFSNESA-BETI 58  
Db 1 MNCKKIFITLISLVSFLPGVSFSDVIOENNPVGSVYISAKYMTPTASHFGKRSIKEDS 60

Qy 59 PGITKKIFAL-----GLDKSEINTHSNFT-RSYDPTYASS-FAGFSGIIGYVNDFRVEF 111  
Db 61 RD-TKAVFLGKKDWDGVKTPSGNTNSITFEKDYSEKYNENPFLGPAAGAVGSMGPRIEP 119

Qy 112 EGSYENFEPERQW--YPENSQYKFFALSRNATNSDNKFIIVLENNGVVDKSLNVNVCYDI 169  
Db 120 EVSYETFDVRNPGGNYKNDHRYCALDHTASSSTAGATTSMVKVKNENLTDISMLNACYDI 179

Qy 170 ASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVNMTLFGGYYHKVGVGD 229  
Db 180 MLDGMPVSPYVCAGIGTDLVSVINATNPKLSYQKGLGISYINPEASVFIGGHFHRVIGN 239

Qy 230 RHERVEIAYHPTALSDVPRITTS-----ASATLNTDYFGWEIGGRF 269  
Db 240 EFKDIATSKVTSSGNASSAVSPGFASAILDVCHFGIELGGRF 282

RESULT 4  
JE0220  
28k surface antigen 1 - Ehrlichia canis  
C:Species: Ehrlichia canis  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0220  
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746  
A:Accession: JE0220  
A:Molecule type: DNA  
A:Residues: 1-287 <RED>  
A:Cross-references: UNIPROT:Q9ZGJ0; UNIPARC:UPI0000034790; GB:AF062762; NID:g3327964; PID

Query Match 26.4%; Score 379; DB 2; Length 287;  
Best Local Similarity 34.1%; Pred. No. 3e-24;  
Matches 101; Conservative 55; Mismatches 102; Indels 38; Gaps 12;

Qy 1 MNYKR-FVVG--VTLSTFVFLSDGAFSDANFSEGRRLG-LYIGSQYKVGIPNFSNESA-BETI 57  
Db 1 MNYKTFITVALVLLTSFTFIP--FVSPARST-IHNFYISGKYMPTASHFGIFSAAKEE 57

Qy 58 IPGITKKIFALGDK-----SEINTH-----SNFRSYDPTYASS-FAGFSGIIGYVNDFR 108  
Db 58 -QSPTKVL--VGLDQRLSHNIINNNTAKSLKVQNSFKYKNNPFLGFAAGIAGYSIGNSR 114

Qy 109 VEFEGSVENFEPERQWYPENSQYKFFALSRNA-----TNSDNKFIIVLENNGV 156  
Db 115 IELEVSHEIFDTKNPGNLYNDSHKYCALSHGSHICSDGNSGDWYTAKTOKFVLLKNEGL 174

Qy 157 VDKSLNVNVCYDIASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVNMT 216  
Db 175 LDVSPMLNACYDITTEKMPSPYICAGIGTDLISMETTONKISYQKGLGNTINSRVS 234

Qy 217 LFGGYYHKVGVGRHERVEIAYHPTAL---SDVPRITTSASATLNTDYFGWEIGGRF 269  
Db 235 VFAGGHFKVIGNEFKGI-----PTLLPDGNSIKVQOSATVTLDVCHFGLEIGSRF 285

RESULT 5  
JE0219  
28k surface antigen 2 - Ehrlichia chaffeensis  
N:Alternate names: MAP1  
C:Species: Ehrlichia chaffeensis  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0219  
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746  
A:Accession: JE0219  
A:Molecule type: DNA  
A:Residues: 1-286 <RED>  
A:Cross-references: UNIPROT:O52105; UNIPARC:UPI000003478C; GB:AF062761

Query Match 26.4%; Score 378.5; DB 2; Length 286;  
Best Local Similarity 31.4%; Pred. No. 3.3e-24;  
Matches 93; Conservative 58; Mismatches 104; Indels 41; Gaps 9;

Qy 1 MNYKRFVGVVTLSTFVFLSDGAFSDANFSEGRRLG-LYIGSQYKVGIPNFSNESA-BETI 59  
Db 1 MNCKKFFITVALTLLMSFLPGISLSDPVDQDINSFNFTVSGKMPFSASHGFMFSKAKEN-R 59

Qy 60 GITKKIFALGDL-----KSEINTHSNFTSYDPTYASS-FAGFSGIIGYVNDFRV 109  
Db 60 NTTVGFIGBQDRCRVCISRTTSLDIFTVPNYSFKYE---NNLFSGFAGIAGYSMDGPR 116

Qy 110 EFGSVENFEPERQWYPENSQYKFFALSR-----NATNSDNKFIIVLENNGVVDKSL 161  
Db 117 ELEVSHEAFDVKNQGNKYKNDHRYALSHLLTGTETQIDGAGSASVPLI--NEGLDKSP 174

Qy 162 NVNVCYDIASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVNMTLFGGG 221

```
Db      175 MLNACYDVISGIPFPSPYICAGIGIDLVSMEFANPKISYQKGLSLSPISPEASVFIGG 234
Qy      222 YHKVVGDRHERVEIAYHPTALSDVPRPTS-----ASATLNTDYFGWEIGFRF 269
Db      235 HPKHVIGNEPRDI-----PTM---IPSSALAGKNYPAIVLDFVFFGIELGGRF 282

RESULT 6
JE0218
28k surface antigen 5 - Ehrlichia chaffensis
N:Alternate names: MAP1
C:Species: Ehrlichia chaffensis
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C:Accession: JE0218
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A:Reference number: JE0216; MUID:98321180; PMID:9647746
A:Accession: JE0218
A:Molecule type: DNA
A:Residues: 1-276 <RED>
A:Cross-references: UNIPROT:Q93DD2; UNIPROT:Q9ACI9; UNIPARC:UPI000003478F; GB:AF062761

Query Match      26.3%; Score 377; DB 2; Length 276;
Best Local Similarity 29.8%; Pred. No. 4.2e-24;
Matches 87; Conservative 57; Mismatches 106; Indels 42; Gaps 6;

Qy      1 MNYKRFVGVGVTSTVFVFLSDGAFSDANFSEGRGLXIGSQYKVGIPNFSNFSABETIPG 60
Db      1 MNYKVFITSALISLISLPGVFSFDGAGNGNFYISGKTPSASHFGVFSABEE-RN 59

Qy      61 ITKKIFAL-----GLDKSEINTHSNFTSRSDPTVYASSFAGSGIIGYVNDFRVE 110
Db      60 TTGVFGLKQNDGSAISNSPNDVTVSNYSFKYE---NNPFLGFAAGAGYSMDGPRIE 116

Qy      111 FPGSYENFEPERQWTPENSQSYKFFALSRNA-----TNSDNKPIVLNNGVVDKSLNVVC 166
Db      117 LEVSYETFDVKQNGNKNYNEAHRYCALSHNSAADSSASNNFVFLNKEGLDLSFMLNAC 176

Qy      167 YDIASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVTMLFGGYYHKV 226
Db      177 YDVVGEGIPFPSPYICAGIGIDLVSMEFANPKISYQKGLSLSPISPEASVFIGHFHKV 236

Qy      227 VGDHREHVEIAYHPTALSDVPRPTSASAT-----INTDYFGWEIG 266
Db      237 IGNE-----FRDIPITIGSTLAGKNYPAIVLDFVCHFGIEMG 276

RESULT 7
I39648
major surface protein 4 - Anaplasma marginale
C:Species: Anaplasma marginale
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C:Accession: I39648
R:Oberle, S.M.; Barbet, A.F.
Gene 136, 291-294, 1993
A:Title: Derivation of the complete msp4 gene sequence of Anaplasma marginale without cl
A:Reference number: I39648; MUID:94124017; PMID:8294020
A:Accession: I39648
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-282 <RES>
A:Cross-references: UNIPROT:Q07408; UNIPARC:UPI0000000B7F; GB:L01987; NID:gi42174; PIDN:
C:Genetics:
A:Gene: msp4

Query Match      19.4%; Score 278; DB 2; Length 282;
Best Local Similarity 26.1%; Pred. No. 8.2e-16;
Matches 77; Conservative 57; Mismatches 119; Indels 42; Gaps 10;

Qy      1 MNYKR-FVGVGVTSTF--VFFLSDGA-----FSDANFSEG---RRGLYIGSQYKVGIPNF 49
```

```
Db      1 MNYRELFTGGLSAATVCCACSLLVSGAVVASPMSEHVASEGVGVMGSPVYVCAAYSPAPFSV 60
Qy      50 SNFSAEETIPGITYKKIPALGLDKS-----EINTHSNFTSRSDPTVYASS---FAGSGIIGY 102
Db      61 TSDMRES---SKETSYVRGYDKSIATIDVSVPAFNSKS-GYTFAPSKNLTISFDGAVGY 116
Qy      103 YVNDPRVFPFGSYENFEPERQWTPENSQSYKFPALSRNATNSDNKFTVLNNGVVDKSLN 162
Db      117 SLGGARVELEASRYRRFATLADGGYAKSAGESLAAITRDNITETNYFVKIDEITNTSVM 176
Qy      163 VNVYDIASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVTMLFGGYY 222
Db      177 LNCYDVLHTDLPSVPVVCAGIGAFVDISKQVTKLAVRGKVGISYQFTPEISLVAGGF 236
Qy      223 YHKVVGDRHERVEIAYHPTALSDVPRPTS-----ASATLNTDYFGWEIGFRF 269
Db      237 YHGLFDESYK-----DIPAHNSVKFSGEAKASVKAHTIADYGFNLGARF 279

RESULT 8
JE0221
28k surface antigen 2 - Ehrlichia canis
C:Species: Ehrlichia canis
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C:Accession: JE0221
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A:Reference number: JE0216; MUID:98321180; PMID:9647746
A:Accession: JE0221
A:Molecule type: DNA
A:Residues: 1-133 <RED>
A:Cross-references: UNIPROT:O85360; UNIPARC:UPI00000034791; GB:AF062762; NID:93327964; PI

Query Match      10.2%; Score 146.5; DB 2; Length 133;
Best Local Similarity 33.6%; Pred. No. 3e-05;
Matches 45; Conservative 23; Mismatches 51; Indels 15; Gaps 7;

Qy      1 MNYKR-FVGVGVTSTVFVFLSDGAFSDANFSEGRRLXIGSQYKVGIPNFSNFSABETI 58
Db      1 MCKKVFITISALISS-IVFLPNVSYSNPVVGNMYGNFYISGKTPSPVPHFGIFSABEE- 58

Qy      59 PGITYKKIPALGLDK-----SEINTHSNFT-RSDPTVYASS-FAGSGIIGYVNDPRV 109
Db      59 --KKKTTVVYGLKENWAGDAISSQSPDNFTIRNYSFKYASNKFLGFAVAIGSIGSPRI 116

Qy      110 BFEGSYENFEPERQ 123
Db      117 EVMSYEAFDVKNQ 130

RESULT 9
G64620
iron(III) dicitrate transport protein - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: G64620
R:Tomb, J.F.; White, O.; Kexlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: G64620
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-787 <TOM>
A:Cross-references: UNIPROT:O25487; UNIPARC:UPI000000D3029; GB:AE000592; GB:AE000511; NID
C:Superfamily: ferrichrome-iron receptor; tonB-dependent receptor amino-terminal homolog
C:Keywords: membrane protein

Query Match      7.1%; Score 102; DB 2; Length 787;
```





Db 249 WYFIGKLSLQDDP---PEKS-----QRMQLAKNGSIYLTATYLRNKKAVDERKFEIK 299  
QY 153 NNG-----VVD-KSLNVNVCY 167  
Db 300 DSGMNAVVIDFKDDNGNLTY 319

RESULT 13  
F82506  
probable TonB system receptor VCA0064 [imported] - Vibrio cholerae (strain N16961 serogroup O139)  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: F82506  
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: F82506  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-713 <HEI>  
A:Cross-references: UNIPROT:Q9KNA1; UNIPARC:UPI00000C33F5; GB:AE004349; GB:AE003853; NID  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA0064  
A:Map position: 2

Query Match 6.8%; Score 97.5; DB 2; Length 713;  
Best Local Similarity 21.2%; Pred. No. 3.5;  
Matches 65; Conservative 36; Mismatches 116; Indels 89; Gaps 14;

QY 21 DGAFTDANFSEGRGLYIGSQYKVGTPNFSN-----PSAETIPGKTK---- 63  
Db 388 DLQADEPAPAR-----ETRFVYGDNLGDLWTLAGLRDPAQKLSPTNDRHG 440

QY 64 -KIFALGLDK-----SEINTNSNFTN-----SYDPTVASSFAGFGIIGYVN 105  
Db 441 YKVVTMGSEWSPSASISYQHPWNTYLSYNGHFRAPSYDAYGASDHSFVPLTPFII- 499

QY 106 DFRVEPEGSYENFEPERQWYKPFALSRNATNSDNKPFVLNNGVVDKSLN-V 164  
Db 500 -----KPNKLAETSDS---FELGSKYDNGQTQFYAVAFYSIFDNFIDVKQ 543

QY 165 VCYDIASGSIPLAPYMCAGVADYIKFLGLSLP-----KFSY-QVKFGVNP 210  
Db 544 VGYDNATGSIQYQYQNIAGV-----KTYGAEMSVMRLLDRWSVENKLGVDGKGENY 598

QY 211 LNVNTMLFGGGY---YHKVVGDRHERVEIAYHPTALSDVPRITS-----ASATLNTDYFGW 263  
Db 599 VRTLTPLEGSVQLNYQERWDAYSRLNWA---SAMSERVTCITEQKETECACTTGWVSM 655

QY 264 BIGFRF 269  
Db 656 DIGLNY 661

RESULT 14  
B71315  
conserved hypothetical protein TP0515 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: B71315  
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: B71315  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-991 <COL>  
A:Cross-references: UNIPROT:O83528; UNIPARC:UPI00000C0A89; GB:AE001227; GB:AE000520; NID  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0515

Query Match 6.8%; Score 97.5; DB 2; Length 991;  
Best Local Similarity 22.0%; Pred. No. 5.4;  
Matches 51; Conservative 36; Mismatches 94; Indels 51; Gaps 9;

QY 42 YKVGIPNFSNPSAETIPGKTKIFALGLDKSEINTNSNFTRSYDPTVASSFAGFGIIG 101  
Db 723 YAYSIAFLSP-----PYISATLSARFLDPQVRAELDFRMPYEITCKQT----- 767

QY 102 YYVNDFRVEPEGSYENFEPERQWYKPFALSRNATNSDNKPFVLNNGVVDKSL 161  
Db 768 -YVYDIGKMSDSY-----EVSGLGWKFSLSYLLKGESGKAAPSGNNGLSISKL 815

QY 162 NVNVCYD-----IASGSIPLAPYMCAGVADYIKFLGLSLPKFSYQVKFGVN--YPLN 212  
Db 816 KLTLSHDDPFLTVRFWKRRIKIQGTLSSELEINF-KLDKSHISFSPDLITTSYIKFLDLS 874

QY 213 VNTMLFGGGYVHKVVGDRHERVEIAYHPTALSDVPRITSASATLNTDYFGWE 264  
Db 875 INTVI-----KNEKL-TPYFPTQSSQSPOTKLWDAFVSSLYP-WD 912

RESULT 15  
B64596  
toxin-like outer membrane protein HP0610 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: B64596  
R:Tomb, J.F.; White, O.; Kexlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, S.; J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.C.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 533-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: B64596  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1943 <TOM>  
A:Cross-references: UNIPROT:O25331; UNIPARC:UPI00000D30AD; GB:AE000575; GB:AE000511; NID  
C:Genetics:  
A:Start codon: GTG

Query Match 6.7%; Score 96.5; DB 2; Length 1943;  
Best Local Similarity 22.9%; Pred. No. 16;  
Matches 62; Conservative 42; Mismatches 102; Indels 65; Gaps 16;

QY 1 MNYKFFVVGVTI-----STFVFFLSDGAFSDANFSEGRRLYIGSOYKVGIPN 48  
Db 85 LNFQATVVGVTLLGSAQKANNNGSIWFGKNLLYLHGNFN--ATNIFLTNFWVGNPN 142

QY 49 FS-----NFSABETI--PGITKPIF---ALGLDKSEINTNS--NFTRSYDPTYA--SSFAG 95  
Db 143 ACGGATINFNADETLNADGLNVTNFQTVAGLQTS-ASQHSWANFNSKLSMEIKNSNFRD 201

QY 96 FS-GIIGYVNDFRVEFE-----GSYENPEPERQWYKPFALSRNATNSDNKPFVLNNGVVDKSLNVCYDIASGSIPLAPYMCAGV-- 134  
Db 202 FTWG--GFNFNSGRITFTFTSGWTNNGATGESGSYVNMVNTDLIFSNILOGGRIY 259

QY 135 FALSBNATNSDNKPFVLNNGVVDKSLNVCYDIASGSIPLAPYMCAGV-- 185  
Db 260 DLKANIIIFNNSQWVIDVSKVNSQSLNGNVTFN--NSRLSVKPNAAINIGDSQQTQTALE 317

QY 186 -ADYIKFLGLISLPKFSYQVKF-GVNYPLNVN 214  
Db 318 NASSLSFYNNVSNVNFNGTTFNGVSY-LNLN 347

Search completed: December 28, 2005, 15:38:31  
Job time : 27 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:38:40 ; Search time 18 Seconds  
(without alignments)  
112.748 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNVRFVGVVTLSTFVFLS.....ASATLNTDYFGWIGRFRAL 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 57103 seqs, 7488799 residues

Total number of hits satisfying chosen parameters: 57103

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_New:  
1: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US03\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	5.9	1565	6	US-10-467-657-2704
2	84	5.9	477	7	US-11-055-822-98
3	83.5	5.8	1433	7	US-11-094-519A-40
4	81	5.6	876	7	US-11-077-550-128
5	81	5.6	877	7	US-11-077-550-126
6	81	5.6	877	7	US-11-077-550-130
7	81	5.6	881	7	US-11-077-550-124
8	81	5.6	902	7	US-11-077-550-132
9	81	5.6	912	7	US-11-077-550-116
10	81	5.6	914	7	US-11-077-550-120
11	81	5.6	944	7	US-11-077-550-122
12	81	5.6	950	7	US-11-077-550-118
13	78	5.4	219	7	US-11-170-653-29
14	77.5	5.4	1122	7	US-11-191-374-3
15	77.5	5.4	1122	7	US-11-191-375-3
16	77.5	5.4	1122	7	US-11-191-588-3
17	76	5.3	1296	6	US-10-615-668-3
18	76	5.3	1468	6	US-10-467-657-1088
19	76	5.3	2053	7	US-11-013-759-9
20	75	5.2	1992	7	US-11-013-759-3
21	75	5.2	1992	7	US-11-013-759-13
22	75	5.2	2047	7	US-11-013-759-4
23	75	5.2	2047	7	US-11-013-759-7
24	74	5.2	826	6	US-10-793-626-1066
25	73	5.1	278	6	US-10-467-657-7386

26	72.5	5.1	265	6	US-10-467-657-1778	Sequence 1778, Ap
27	72	5.0	475	6	US-10-510-386-98	Sequence 98, Appl
28	71.5	5.0	173	6	US-10-984-376-13	Sequence 13, Appl
29	71.5	5.0	408	6	US-10-793-626-2286	Sequence 2286, Ap
30	71.5	5.0	527	7	US-11-120-543-18	Sequence 18, Appl
31	71	5.0	497	7	US-11-037-829A-15	Sequence 15, Appl
32	70.5	4.9	174	6	US-10-984-376-10	Sequence 10, Appl
33	70	4.9	222	7	US-11-170-653-32	Sequence 32, Appl
34	70	4.9	223	7	US-11-170-653-30	Sequence 30, Appl
35	70	4.9	223	7	US-11-170-653-31	Sequence 31, Appl
36	69.5	4.8	174	6	US-10-984-376-7	Sequence 7, Appl
37	69.5	4.8	174	6	US-10-984-376-8	Sequence 8, Appl
38	69.5	4.8	221	7	US-11-170-653-44	Sequence 44, Appl
39	69.5	4.8	428	7	US-11-074-176-364	Sequence 364, App
40	69	4.8	175	6	US-10-467-657-2512	Sequence 2512, Ap
41	69	4.8	190	7	US-11-170-653-33	Sequence 33, Appl
42	69	4.8	381	6	US-10-793-626-184	Sequence 184, Appl
43	68.5	4.8	174	7	US-11-103-957-90	Sequence 90, Appl
44	68.5	4.8	1259	6	US-10-467-657-5510	Sequence 5510, Ap
45	68.5	4.8	1438	6	US-10-511-559-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1  
US-10-467-657-2704  
; Sequence 2704, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; PRIOR FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: Seqwin99, version 1.04  
; SEQ ID NO 2704  
; LENGTH: 1565  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-2704

Query Match	5.9%	Score 85;	DB 6;	Length 1565;
Best Local Similarity	21.7%	Pred. No. 3.9;		
Matches	60;	Conservative	37;	Mismatches 96; Indels 84; Gaps 15;
Qy	2	NYKRFVGVGT-LSTFVFFLSDGAFSDANFSEGRGLVIGSOY-KVGIPNFSNFSAEETIP	59	
Db	17	NKGRFFVGATDLV-----KNKRQGNIGNALSVPMDFSVADNKRKA	60	
Qy	60	GITKKIFALGLDKSEINTHSNFTSRSDPTVASSPAGFSGIIGYVNDPFRVEFGSYENFE	119	
Db	61	TVDPQYAVSVKKAKEVHT-----FYYQVNGHNDVAD-KENEYRVV---EQNNYE	108	
Qy	120	PEROWYPENSQSYKFFALSRSNATNSDNKFLVLENNGVVDKSLNVNVCYOTASGSIPLAPY	179	
Db	109	PHKAWASNLGRLEDYNNAR-----FNKFVT-----EVA-----PIAP-	141	
Qy	180	MCAGVGADYIKFLGISLPKFSYQVKGNYVPLNVNTMLFGGGYVHKVVDGHERVEAYH	239	
Db	142	TDAGGGJDTYK-----DKNRFSFVRIGAG-----RQLVYKGVYHQ-----EGNEKGYD	186	
Qy	240	PTALSDVPRTTTASA-----TLNTDYFGWEIFG	267	
Db	187	LRDLQAYRYAIACTPYKDNIDQTMTEGL---IGF	220	

```
RESULT 2
US-11-055-822-98
; Sequence 98, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogar, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 98
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-98

Query Match      5.9%; Score 84; DB 7; Length 477;
Best Local Similarity 20.9%; Pred. No. 1;
Matches 56; Conservative 32; Mismatches 82; Indels 98; Gaps 15;

QY 11 TLSTFVPFLSDG--AFSDANFSGRGLYIGSVKVGIP--NFSNFSABE-----TIP 59
DB 5 TPBEIVKFDENVEFVDVFTD-----LPGTEQHFSIPAASFDADTIEGLAFDGSIR 59
QY 60 GITKIPFALGLKSEIN-----THSNFTR---SYDP-----87
DB 60 GFTT-----IDSDMNLPLDGLTATLDPFRKAKTLNVKFFVHDPTFREAFSRDPRNVAR 113
QY 88 -----TVASSFAGSGIIGYVND---FRVEFGSYENFEPERQWYPENSQSYKFFA 136
DB 114 KAEQYLASTGIADTCNFGABAEFLYFDSVRYSTEMNSGFYEVDTEBGMNRGKE-----167
QY 137 LSRNATNSDNKFVLENNGVVDSKSLNVNVCYDITASGSIPLAPY-MCAGVGADYIKFL---192
DB 168 -----TN-----LDGTPNLGAKNVKGGYFFVAPYDQTVDDVMVRNLAS 209
QY 193 GISLPKFSYQVKFG-----VNYPLNVNTML 217
DB 210 GFALERPHHEVGGQOBEIN--RFNTML 235

RESULT 3
US-11-094-519A-40
; Sequence 40, Application US/11094519A
; Publication No. US20050281810A1
```

```
; GENERAL INFORMATION:
; APPLICANT: BERNSTEIN, Jeanne
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0140P
; CURRENT APPLICATION NUMBER: US/11/094,519A
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: US/09/695,293
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: IL 132558
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-094-519A-40

Query Match      5.8%; Score 83.5; DB 7; Length 1433;
Best Local Similarity 20.5%; Pred. No. 4.8; Mismatches 40; Indels 65; Gaps 12;
Matches 54; Conservative 40;

QY 39 GSQYKV-GIPNFSNFSABETI-----PGI-TKKIFALGLD-----KSEINTHSNF 81
DB 581 GLSYKLEGLKKFTEYSLRFLAYNRYPGVSTDDITVTLSDVPSAPPQNVSLVNSRSI 640
QY 82 TRSYDPTVASSPAGSGIIGYVNDFRVEFGSYENFEPERQWY-----PENSO-SYKFPF 135
DB 641 KVSMLPPPSGTQNGF--ITGVKIRHKTRRGEMETLEPNLWYFTGLEKGSQYSFQVS 698
QY 136 ALSRNATNSDNKFIV-----LENNGVVDSKSLNVN-----CYDIASGSIPLAPYMC- 182
DB 699 AMTVNGTGPSSNWNYPATPENDLDESQVDPQSSLHVPRPQTNCI-IMSWTPELNINIVVR 757
QY 183 -----GVGADYIKFLGSLPKFSYQVKFGVNVYPLNVNTMLPGGGYHYHKVVGDRHERVEI 236
DB 758 GYIIGVGSGPYAEIVRVDSKQRYYSIE-----RLESSSHYVISL 797
QY 237 AYHPTALSDVPRPTTSASATLNTD 259
DB 798 KAFNNAGEGVPLYESATTRSITD 820

RESULT 4
US-11-077-550-128
; Sequence 128, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
```



Db 208 RFM--LTYSNATNDVGBGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 257  
Qy 56 ETIPGITKKIF---ALGLDKSEINTHSNFTSRSDPTYASSF----- 93  
Db 258 QTISSTVSNIFYSQYNVKLSEYAEIYAFGGPTIDILPKSARKYFEKALDYRSIAKRLNS 317  
Qy 94 -----AGFSGIIGYY----VNDPRVFEFGSYE---NFEPERQWYPENSQSYKFFALSRLN 140  
Db 318 ITTANPSSFNKYIGEYKQKLRKYRFVVBSSGGEVTVNRNKFVELYNELTQIFTEFNAYK- 376  
Qy 141 ATNSDNKFIIVLEN--NGVVDKSLNVNVCYDIAGSIP LAPYMCAGGADYIKFLGSLPK 198  
Db 377 IYNVQNRKIYLSNVYTPVTANILDDNV-YDION-----GNIPK 414  
Qy 199 PSYQVKF-GVNY-----PLNVNTMLF-----GGYYHKVVG----- 228  
Db 415 SNLNVLFMGQNLNRNPALRKVNPNENMLYFTKFKHKAIDGRSLYNTKLDCRELLVKNKTDL 474  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 475 PFIGDISDKTDFILRKDINEETEVYYPDNVS 507

## RESULT 7

US-11-077-550-124  
; Sequence 124, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 124  
; LENGTH: 881  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-124

Query Match 5.6%; Score 81; DB 7; Length 881;  
Best Local Similarity 20.4%; Pred. No. 4.4;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVVGVTLSFTVFFLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSFSAE 55  
Db 212 RFM--LTYSNATNDVGBGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 261  
Qy 56 ETIPGITKKIF---ALGLDKSEINTHSNFTSRSDPTYASSF----- 93  
Db 262 QTISSTVSNIFYSQYNVKLSEYAEIYAFGGPTIDILPKSARKYFEKALDYRSIAKRLNS 321  
Qy 94 -----AGFSGIIGYY----VNDPRVFEFGSYE---NFEPERQWYPENSQSYKFFALSRLN 140

Db 322 ITTANPSSFNKYIGEYKQKLRKYRFVVBSSGGEVTVNRNKFVELYNELTQIFTEFNAYK- 380  
Qy 141 ATNSDNKFIIVLEN--NGVVDKSLNVNVCYDIAGSIP LAPYMCAGGADYIKFLGSLPK 198  
Db 381 IYNVQNRKIYLSNVYTPVTANILDDNV-YDION-----GNIPK 418  
Qy 199 PSYQVKF-GVNY-----PLNVNTMLF-----GGYYHKVVG----- 228  
Db 419 SNLNVLFMGQNLNRNPALRKVNPNENMLYFTKFKHKAIDGRSLYNTKLDCRELLVKNKTDL 478  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 479 PFIGDISDKTDFILRKDINEETEVYYPDNVS 511

## RESULT 8

US-11-077-550-132  
; Sequence 132, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 132  
; LENGTH: 902  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-132

Query Match 5.6%; Score 81; DB 7; Length 902;  
Best Local Similarity 20.4%; Pred. No. 4.6;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVVGVTLSFTVFFLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSFSAE 55  
Db 212 RFM--LTYSNATNDVGBGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 261  
Qy 56 ETIPGITKKIF---ALGLDKSEINTHSNFTSRSDPTYASSF----- 93  
Db 262 QTISSTVSNIFYSQYNVKLSEYAEIYAFGGPTIDILPKSARKYFEKALDYRSIAKRLNS 321  
Qy 94 -----AGFSGIIGYY----VNDPRVFEFGSYE---NFEPERQWYPENSQSYKFFALSRLN 140  
Db 322 ITTANPSSFNKYIGEYKQKLRKYRFVVBSSGGEVTVNRNKFVELYNELTQIFTEFNAYK- 380  
Qy 141 ATNSDNKFIIVLEN--NGVVDKSLNVNVCYDIAGSIP LAPYMCAGGADYIKFLGSLPK 198  
Db 381 IYNVQNRKIYLSNVYTPVTANILDDNV-YDION-----GNIPK 418

Qy 199 PSYQVKP-GVNY-----PLNVNTMLF-----GGGYHKVVG----- 228  
Db 419 SNLNVLMQNLNRPNALRVKPNMMLYLFTKFCHKAIDGRSLYNTKLDCRELLVKNTDL 478  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 479 PFIGDISDKVTDIFLRKDINEETEVIYYPDNVS 511

## RESULT 9

US-11-077-550-116  
; Sequence 116, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 116  
; LENGTH: 912  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-116

Query Match 5.6%; Score 81; DB 7; Length 912;  
Best Local Similarity 20.4%; Pred. No. 4.6;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVVGVTLSFTVFFLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSNFSAE 55  
Db 212 RFM--LTYSNATNDVGEGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 261  
Qy 56 ETIPGITKKIF-----ALGLDKSEINTHSNFTSRSDPTVASSF----- 93  
Db 262 QTISSTVSNIFYSQYNVKLEAYEIAFGGPTIDLPKSARKYFEEKALDYRSTAKRLNS 321  
Qy 94 -----AGFSGIIGYV-----VNDPVEFEGSYE---NFEPEROWYPENSQYKFPALSRN 140  
Db 322 ITTANPSSFNKYIGYKQKLIRKYRFVVESSGEVTVNRNKFVELYNELTQIFTEFNAYAK- 380  
Qy 141 ATNSDNKFTVLEN--NGVVDKSLNVNVCYDIAGSISPLAPYMCAGVADYIKFGLISLPK 198  
Db 381 IYVQNRKYLNSVYTPVTANILDDNV-YDIQN-----GFNIPK 418  
Qy 199 FSQVQKP-GVNY-----PLNVNTMLF-----GGGYHKVVG----- 228  
Db 419 SNLNVLMQNLNRPNALRVKPNMMLYLFTKFCHKAIDGRSLYNTKLDCRELLVKNTDL 478  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 479 PFIGDISDKVTDIFLRKDINEETEVIYYPDNVS 511

## RESULT 10

US-11-077-550-120  
; Sequence 120, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 120  
; LENGTH: 914  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-120

Query Match 5.6%; Score 81; DB 7; Length 914;  
Best Local Similarity 20.4%; Pred. No. 4.6;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVVGVTLSFTVFFLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSNFSAE 55  
Db 212 RFM--LTYSNATNDVGEGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 261  
Qy 56 ETIPGITKKIF-----ALGLDKSEINTHSNFTSRSDPTVASSF----- 93  
Db 262 QTISSTVSNIFYSQYNVKLEAYEIAFGGPTIDLPKSARKYFEEKALDYRSTAKRLNS 321  
Qy 94 -----AGFSGIIGYV-----VNDPVEFEGSYE---NFEPEROWYPENSQYKFPALSRN 140  
Db 322 ITTANPSSFNKYIGYKQKLIRKYRFVVESSGEVTVNRNKFVELYNELTQIFTEFNAYAK- 380  
Qy 141 ATNSDNKFTVLEN--NGVVDKSLNVNVCYDIAGSISPLAPYMCAGVADYIKFGLISLPK 198  
Db 381 IYVQNRKYLNSVYTPVTANILDDNV-YDIQN-----GFNIPK 418  
Qy 199 FSQVQKP-GVNY-----PLNVNTMLF-----GGGYHKVVG----- 228  
Db 419 SNLNVLMQNLNRPNALRVKPNMMLYLFTKFCHKAIDGRSLYNTKLDCRELLVKNTDL 478  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 479 PFIGDISDKVTDIFLRKDINEETEVIYYPDNVS 511

## RESULT 11

US-11-077-550-122  
; Sequence 122, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:



```

/ APPLICANT: Wayne, Jonathan
/ TITLE OF INVENTION: Recombinant Toxin Fragments
/ FILE REFERENCE: 1581.0130004
/ CURRENT APPLICATION NUMBER: US/11/077,550
/ CURRENT FILING DATE: 2005-03-11
/ PRIOR APPLICATION NUMBER: 10/241,596
/ PRIOR FILING DATE: 2002-09-12
/ PRIOR APPLICATION NUMBER: 09/255,829
/ PRIOR FILING DATE: 1999-02-23
/ PRIOR APPLICATION NUMBER: PCT/GB97/02273
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 08/782,893
/ PRIOR FILING DATE: 1996-12-27
/ PRIOR APPLICATION NUMBER: GB9625996.5
/ PRIOR FILING DATE: 1996-12-13
/ PRIOR APPLICATION NUMBER: GB9617671.4
/ PRIOR FILING DATE: 1996-08-23
/ NUMBER OF SEQ ID NOS: 179
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 118
/ LENGTH: 950
/ TYPE: PRT
/ ORGANISM: Clostridium botulinum
US-11-077-550-118

Query Match          5.6%; Score 81; DB 7; Length 950;
Best Local Similarity 20.4%; Pred. No. 4.9;
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 161;

Qy  5 RFVGVTLSTFVFLSDGASDANP-----SEGRRLGYTGSQYKVGIPNFSNFAE 55
Db  212 RFM--LTYSNATNDVGEGRFSSEFCMDPILILMHELNHAMH--NLXGIAIPN-----D 261
Qy  56 ETIPGTTKTF-----ALGDKSEINTHSNTRSVDPTYASSF----- 93
Db  262 QTSSVTSNIFYSQYNVKLEYABIAFPGPTDILPKSARKYFEKALDYVRSIAKLNS 321
Qy  94 -----AGFGIIGIY-----VNDPRVFEFGSYE---NFEPEROWYPENSQSYKPFALSRN 140
Db  322 ITTANPSSFNKYIGEYKQKLIRKYRVFVSSGEVTYNNKFEVELYNELTQIFTEFNyak- 380
Qy  141 ATNSDNKFIVLN--NGVVDKSLNVNVCYDIAGSIPLAPYMCAGVADYIKFISLPLK 198
Db  381 IYNVQRKLYLSNVYTPVTANILDDNV-YDION-----GFIPIK 418
Qy  199 FSQVQKF-GVNY-----PLNVNTMLF-----GGYVHKVVG----- 228
Db  419 SNLNVLFPMQNLGRNPAIRKVPENMLYLFTKFCHKAIDGRSLYNKTLDCRELLVKNITDL 478
Qy  229 -----DRHERVEIAHYHPTALS 244
Db  479 PFIGDISDVKTDFLRKDKINEETEVYYPDNVS 511

RESULT 13
US-11-170-653-29
/ Sequence 29, Application US/11170653
/ Publication No. US2005027169A1
/ GENERAL INFORMATION:
/ APPLICANT: Danisco A/S
/ APPLICANT: Sibbesen, Ole
/ APPLICANT: Sorensen, Jens
/ TITLE OF INVENTION: Xylanase Variants Having Altered Sensitivity to Xylanase
/ FILE REFERENCE: 674509-2046
/ CURRENT APPLICATION NUMBER: US/11/170,653
/ CURRENT FILING DATE: 2005-06-23
/ PRIOR APPLICATION NUMBER: US/10/237,386
/ PRIOR FILING DATE: 2002-09-09
/ PRIOR APPLICATION NUMBER: PCT/IB01/00426
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: GB 0005585.5
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: GB 0015751.1

```

;; PRIOR FILING DATE: 2000-06-27  
;; NUMBER OF SEQ ID NOS: 66  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 29  
;; LENGTH: 219  
;; TYPE: PRT  
;; ORGANISM: C. gracile  
US-11-170-653-29

Query Match 5.4%; Score 78; DB 7; Length 219;  
Best Local Similarity 20.5%; Pred. No. 1.3;  
Matches 43; Conservative 32; Mismatches 75; Indels 60; Gaps 11;  
Qy 9 GVTLSSTVFVLDGAFDANFSEGRGLYIGSQYKVGIPNPNF-SAEETIPGTTKIPFA 67  
Db 38 GTNNGYFYFWTDGG-GTVNYQNG-----AGGSYSVQWQNGFVGGKWNPGGAARTI-- 89  
Qy 68 LGLDKSEINTHSNTRSDPTVYASSPAGFSG-----IIGYVNDFRVEFEFGSYENFEPER 122  
Db 90 -----NFSGTSP-QGNGYLAIYGTQNPVVEYI-----VESFGTYD----- 126  
Qy 123 QWYPENSQYKFFALSNNATNSDNKFIVLNNGVVDKSLNVNVCYD-----IASGS 173  
Db 127 -----PSSQASKEGFIQDQGST-----YTTAKTRVNPQSIETSTFDQFWSVRQNHRSQS 178  
Qy 174 IPLAPYMCAGVADYIKFLGSLPKFSYQV 203  
Db 179 VVAAHFAWAQA-----GLKLGSHNYQI 202

## RESULT 14

US-11-191-374-3  
;; Sequence 3, Application US/11191374  
;; Publication No. US20050260673A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Hresko, Michelle Coutu  
;; APPLICANT: McLaire, Merry B.  
;; APPLICANT: Williams, Deryck J.  
;; APPLICANT: Frevert, Anita M.  
;; APPLICANT: Chiapelli, Brandi  
;; APPLICANT: Baublite, Catherine  
;; APPLICANT: Kloeck, Andrew P.  
;; APPLICANT: Davila-Aponte, Jennifer A.  
;; APPLICANT: Bradley, John D.  
;; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE  
;; FILE REFERENCE: 12557-015001  
;; CURRENT APPLICATION NUMBER: US/11/191,374  
;; PRIOR FILING DATE: 2005-07-28  
;; PRIOR APPLICATION NUMBER: US/10/771,708  
;; PRIOR FILING DATE: 2004-02-04  
;; PRIOR APPLICATION NUMBER: US 60/444,771  
;; NUMBER OF SEQ ID NOS: 54  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 1122  
;; TYPE: PRT  
;; ORGANISM: Strongyloides stercoralis  
US-11-191-374-3

Query Match 5.4%; Score 77.5; DB 7; Length 1122;  
Best Local Similarity 22.6%; Pred. No. 13;  
Matches 36; Conservative 31; Mismatches 57; Indels 35; Gaps 8;  
Qy 118 FEPERQWYPENSQYKFFALS-R-NATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPL 176  
Db 631 YEPEKKYHPKSGKNETS YETGTVNKSNEVESETLTNSGVESGLEKNII--TAPPSIPK 688  
Qy 177 APYMCAG-----VGADYIKFLGSLPKFSYQ-VKFGVNYPLNVNTMLFGGGYHKVVG 228  
Db 689 IP---EGPLVPILIPADQVQTI-----CDYEGIKVQIKSPQSTGTGIVFVKNHYETC-- 737

Qy 229 DRHERVEIAYH-----PTALS DVPRRTTSASATLNT 258  
Db 738 ----RVEVSNSDAATLELGLPASFGMKPVTLSATSSDST 772

## RESULT 15

US-11-191-375-3  
;; Sequence 3, Application US/11191375  
;; Publication No. US20050260674A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Hresko, Michelle Coutu  
;; APPLICANT: McLaire, Merry B.  
;; APPLICANT: Williams, Deryck J.  
;; APPLICANT: Frevert, Anita M.  
;; APPLICANT: Chiapelli, Brandi  
;; APPLICANT: Baublite, Catherine  
;; APPLICANT: Kloeck, Andrew P.  
;; APPLICANT: Davila-Aponte, Jennifer A.  
;; APPLICANT: Bradley, John D.  
;; APPLICANT: Xu, Siqun  
;; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE  
;; FILE REFERENCE: 12557-015001  
;; CURRENT APPLICATION NUMBER: US/11/191,375  
;; CURRENT FILING DATE: 2005-07-28  
;; PRIOR APPLICATION NUMBER: US/10/771,708  
;; PRIOR FILING DATE: 2004-02-04  
;; PRIOR APPLICATION NUMBER: US 60/444,771  
;; NUMBER OF SEQ ID NOS: 54  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 1122  
;; TYPE: PRT  
;; ORGANISM: Strongyloides stercoralis  
US-11-191-375-3

Query Match 5.4%; Score 77.5; DB 7; Length 1122;  
Best Local Similarity 22.6%; Pred. No. 13;  
Matches 36; Conservative 31; Mismatches 57; Indels 35; Gaps 8;  
Qy 118 FEPERQWYPENSQYKFFALS-R-NATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPL 176  
Db 631 YEPEKKYHPKSGKNETS YETGTVNKSNEVESETLTNSGVESGLEKNII--TAPPSIPK 688  
Qy 177 APYMCAG-----VGADYIKFLGSLPKFSYQ-VKFGVNYPLNVNTMLFGGGYHKVVG 228  
Db 689 IP---EGPLVPILIPADQVQTI-----CDYEGIKVQIKSPQSTGTGIVFVKNHYETC-- 737  
Qy 229 DRHERVEIAYH-----PTALS DVPRRTTSASATLNT 258  
Db 738 ----RVEVSNSDAATLELGLPASFGMKPVTLSATSSDST 772

Search completed: December 28, 2005, 15:45:23  
Job time : 19 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:38:05 ; Search time 79 Seconds  
(without alignments)

1433.313 Million cell updates/sec

Title: US-10-731-554-46

Perfect score: 1434

Sequence: 1 MNYKRFVVGVTLSFVFVFLS.....ASATLNTDYFGWIGRFRAL 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1434	100.0	271	3	US-09-811-007-46 Sequence 46, Appl
2	1434	100.0	271	4	US-10-062-624-46 Sequence 46, Appl
3	1434	100.0	271	4	US-10-062-051-46 Sequence 46, Appl
4	1434	100.0	271	4	US-10-062-920-46 Sequence 46, Appl
5	1434	100.0	271	4	US-10-062-920-46 Sequence 46, Appl
6	1434	100.0	271	5	US-10-731-554-46 Sequence 46, Appl
7	1194	83.3	271	3	US-09-846-808-20 Sequence 20, Appl
8	1194	83.3	271	4	US-10-284-986-20 Sequence 20, Appl
9	1194	83.3	271	4	US-10-369-293-20 Sequence 20, Appl
10	1194	83.3	271	4	US-10-285-042-20 Sequence 20, Appl
11	509.5	35.5	280	3	US-09-811-007-42 Sequence 42, Appl
12	509.5	35.5	280	4	US-10-062-624-42 Sequence 42, Appl
13	509.5	35.5	280	4	US-10-059-964-48 Sequence 48, Appl
14	509.5	35.5	280	4	US-10-062-051-42 Sequence 42, Appl
15	509.5	35.5	280	4	US-10-062-920-42 Sequence 42, Appl
16	509.5	35.5	280	4	US-10-314-639-48 Sequence 48, Appl
17	509.5	35.5	280	4	US-10-062-624-42 Sequence 42, Appl
18	509.5	35.5	280	5	US-10-731-554-42 Sequence 42, Appl
19	509.5	35.5	280	5	US-10-901-714-48 Sequence 48, Appl
20	509.5	35.5	280	5	US-10-901-774-48 Sequence 48, Appl
21	507	35.4	283	3	US-09-846-808-14 Sequence 14, Appl
22	507	35.4	283	3	US-09-811-007-10 Sequence 10, Appl
23	507	35.4	283	4	US-10-062-624-10 Sequence 10, Appl
24	507	35.4	283	4	US-10-059-964-4 Sequence 4, Appl
25	507	35.4	283	4	US-10-062-051-10 Sequence 10, Appl
26	507	35.4	283	4	US-10-284-986-14 Sequence 14, Appl
27	507	35.4	283	4	US-10-062-920-10 Sequence 10, Appl

28	507	35.4	283	4	US-10-314-639-4	Sequence 4, Appl
29	507	35.4	283	4	US-10-369-293-14	Sequence 14, Appl
30	507	35.4	283	4	US-10-285-042-14	Sequence 14, Appl
31	507	35.4	283	4	US-10-680-349-10	Sequence 10, Appl
32	507	35.4	283	5	US-10-731-554-10	Sequence 10, Appl
33	507	35.4	283	5	US-10-901-714-4	Sequence 4, Appl
34	507	35.4	283	5	US-10-901-774-4	Sequence 4, Appl
35	494	34.4	285	3	US-09-846-808-9	Sequence 9, Appl
36	494	34.4	285	4	US-10-284-986-9	Sequence 9, Appl
37	494	34.4	285	4	US-10-369-293-9	Sequence 9, Appl
38	494	34.4	285	4	US-10-285-042-9	Sequence 9, Appl
39	492	34.3	285	4	US-10-059-964-30	Sequence 30, Appl
40	492	34.3	285	4	US-10-314-639-30	Sequence 30, Appl
41	492	34.3	285	5	US-10-901-714-30	Sequence 30, Appl
42	492	34.3	285	5	US-10-901-774-30	Sequence 30, Appl
43	489	34.1	279	4	US-10-059-964-24	Sequence 24, Appl
44	489	34.1	279	4	US-10-314-639-24	Sequence 24, Appl
45	489	34.1	279	5	US-10-901-714-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-09-811-007-46  
; Sequence 46, Application US/09811007  
; Publication No. US20030185849A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/811.007  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-09-811-007-46

Query Match	100.0%	Score 1434	DB 3	Length 271
Best Local Similarity	100.0%	Pred. No. 2.2e-134	Mismatches 0	Indels 0
Matches 271	Conservative 0			Gaps 0
Qy	1	MNYKRFVVGVTLSFVFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG	60	
Db	1	MNYKRFVVGVTLSFVFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG	60	
Qy	61	ITKKIFALGDKSINTHSNPTRSYDPTYASSPAGFSGLIIGYVNDFRVEFGSYENPEP	120	
Db	61	ITKKIFALGDKSINTHSNPTRSYDPTYASSPAGFSGLIIGYVNDFRVEFGSYENPEP	120	
Qy	121	ERQWYPENSQSKFALSBNATNDKFI VLENNGVVDKSLNNVVCYDIAGSGIPLAPYM	180	
Db	121	ERQWYPENSQSKFALSBNATNDKFI VLENNGVVDKSLNNVVCYDIAGSGIPLAPYM	180	
Qy	181	CAGVGADYIKFLGILSLPKFSYQVNFVNNVNTMLFGGGYHKVYVGDHRHVEIAVHP	240	
Db	181	CAGVGADYIKFLGILSLPKFSYQVNFVNNVNTMLFGGGYHKVYVGDHRHVEIAVHP	240	
Qy	241	TALSDVPRTTTSASATLNTDYFGWIGRFRAL	271	
Db	241	TALSDVPRTTTSASATLNTDYFGWIGRFRAL	271	

RESULT 2

US-10-062-624-46  
; Sequence 46, Application US/10062624  
; Publication No. US20020115840A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2/DI  
; CURRENT APPLICATION NUMBER: US/10/062,624  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-10-062-624-46  
  
Query Match 100.0%; Score 1434; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.2e-134;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
DB 1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
  
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGFSGIIGYVNDPFRVEFGSYENFEP 120  
DB 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGFSGIIGYVNDPFRVEFGSYENFEP 120  
  
QY 121 ERQWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180  
DB 121 ERQWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180  
  
QY 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHVERVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHVERVEIAYHP 240  
  
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271  
DB 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

RESULT 3  
US-10-062-051-46  
; Sequence 46, Application US/10062051  
; Publication No. US20030073095A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/10/062,051  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US/09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-10-062-051-46

Query Match 100.0%; Score 1434; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.2e-134;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
DB 1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
  
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGFSGIIGYVNDPFRVEFGSYENFEP 120  
DB 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGFSGIIGYVNDPFRVEFGSYENFEP 120  
  
QY 121 ERQWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180  
DB 121 ERQWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180  
  
QY 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHVERVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHVERVEIAYHP 240  
  
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271  
DB 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

RESULT 4  
US-10-062-920-46  
; Sequence 46, Application US/10062920  
; Publication No. US20030096250A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/10/062,920  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US/09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-10-062-920-46  
  
Query Match 100.0%; Score 1434; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.2e-134;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
DB 1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
  
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGFSGIIGYVNDPFRVEFGSYENFEP 120  
DB 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGFSGIIGYVNDPFRVEFGSYENFEP 120  
  
QY 121 ERQWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180  
DB 121 ERQWYPENSQSYKFFALSRNATNSDNKFI VLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180  
  
QY 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHVERVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHVERVEIAYHP 240  
  
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

```
Db      241 TALSDVPRRTTSASATLNTDYFGWEIGFRFAL 271
|||||
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-10-731-554-46

Query Match      100.0%; Score 1434; DB 5; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Db      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Qy      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFEGSYENFEP 120
Db      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFEGSYENFEP 120
Qy      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Qy      181 CAGVGADYIKFLGISLPKFSYQVKGYNVPLNVTMLFGGGYHKVVGDRHERVEIAYHP 240
Db      181 CAGVGADYIKFLGISLPKFSYQVKGYNVPLNVTMLFGGGYHKVVGDRHERVEIAYHP 240
Qy      241 TALSDVPRRTTSASATLNTDYFGWEIGFRFAL 271
Db      241 TALSDVPRRTTSASATLNTDYFGWEIGFRFAL 271

RESULT 7
US-09-846-808-20
; Sequence 20, Application US/09846808
; Patent No. US20020064531A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: Yu, Xu-Jie
; TITLE OF INVENTION: Ehrlichia chaffeensis 28 kDa Outer Membrane
; FILE REFERENCE: D6311
; CURRENT APPLICATION NUMBER: US/09/846,808
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,035
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 20
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: P28-20 Outer Membrane Protein of
; OTHER INFORMATION: Ehrlichia chaffeensis
US-09-846-808-20

Query Match      83.3%; Score 1194; DB 3; Length 271;
Best Local Similarity 78.6%; Pred. No. 2e-110;
Matches 213; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

Qy      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Db      1 MNYKRFVVGVALATLLSFLPDSNFSNFSANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Qy      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFEGSYENFEP 120
Db      61 LTKSIFALGLDKSISDHAGTQAYNPTYASNFAGSGIIGYYVNDPRVFEFEGSYENFEP 120
Qy      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db      121 ERQWYPENGGESHKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
```

---

```
Db      241 TALSDVPRRTTSASATLNTDYFGWEIGFRFAL 271
|||||
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-10-680-349-46

Query Match      100.0%; Score 1434; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Db      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Qy      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFEGSYENFEP 120
Db      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFEGSYENFEP 120
Qy      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Qy      181 CAGVGADYIKFLGISLPKFSYQVKGYNVPLNVTMLFGGGYHKVVGDRHERVEIAYHP 240
Db      181 CAGVGADYIKFLGISLPKFSYQVKGYNVPLNVTMLFGGGYHKVVGDRHERVEIAYHP 240
Qy      241 TALSDVPRRTTSASATLNTDYFGWEIGFRFAL 271
Db      241 TALSDVPRRTTSASATLNTDYFGWEIGFRFAL 271

RESULT 6
US-10-731-554-46
; Sequence 46, Application US/10731554
; Publication No. US20040247616A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xu-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/731,554
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/811,007
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
```





QY 181 CAGVGADYIKFLGSLPKFSYQVKGNYPLNVNTMLFGGYYHKVGDHRRHVEIAYHP 240  
DB 181 CAGVGADYIKFLGSLPKFSYQVKGNYPLNVNTMLFGGYYHKVGNRYERVEIAYHP 240  
QY 241 TALSDVPRTTASATLNTDYGWEIGRPFAL 271  
DB 241 ATLTVNPKTTASATLNTDYGWEIGRPFAL 271

RESULT 11  
US-09-811-007-42  
; Sequence 42, Application US/09811007  
; Publication No. US20030185849A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/811,007  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 42  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein  
US-09-811-007-42

Query Match 35.5%; Score 509.5; DB 3; Length 280;  
Best Local Similarity 40.5%; Pred. No. 4.1e-42;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
QY 1 MNYKRFVVGVTLSFVFFLSDGAFSD---ANPSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLMSILPYQSPADPVGSRNDNKEGFYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIFALGLDKSEINTHNSNFTN-----SYDPTVASSFAGSGIIGYVNDFR 108  
DB 61 PINGTNSLTKKVFLGKKD-GDITKDDFTRVAPGIDFQNNLIS---GFSSGIGYSMDGPR 116  
QY 109 VEPGSGYENFEPEP--QWYPENSQSYKFFALSARNATNSDNKFIVLNNGVVDKSLNVNC 166  
DB 117 IELEAAVQOFPKNTDNDTNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLMVNTC 176  
QY 167 YDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKGNYPLNVNTMLFGGYYHK 225  
DB 177 YDITAGVSFVPYACAGIGADLITIFKDLNL-KFAYQKGIGSYDITEVSAFISAFIGYYHG 235  
QY 226 VVGDHRRHVEIAYHPTALSDVPRTTASATLNTDYGWEIGRPF 269  
DB 236 VIGNKFEKIPV-ITPVVLNDAPQTTASATLNTDYGWEIGRPF 278

RESULT 12  
US-10-062-624-42  
; Sequence 42, Application US/10062624  
; Publication No. US20020115840A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2/D1  
; CURRENT APPLICATION NUMBER: US/10/062,624  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 09/660,587

; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 42  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein  
US-10-062-624-42

Query Match 35.5%; Score 509.5; DB 4; Length 280;  
Best Local Similarity 40.5%; Pred. No. 4.1e-42;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
QY 1 MNYKRFVVGVTLSFVFFLSDGAFSD---ANPSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLMSILPYQSPADPVGSRNDNKEGFYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIFALGLDKSEINTHNSNFTN-----SYDPTVASSFAGSGIIGYVNDFR 108  
DB 61 PINGTNSLTKKVFLGKKD-GDITKDDFTRVAPGIDFQNNLIS---GFSSGIGYSMDGPR 116  
QY 109 VEPGSGYENFEPEP--QWYPENSQSYKFFALSARNATNSDNKFIVLNNGVVDKSLNVNC 166  
DB 117 IELEAAVQOFPKNTDNDTNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLMVNTC 176  
QY 167 YDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKGNYPLNVNTMLFGGYYHK 225  
DB 177 YDITAGVSFVPYACAGIGADLITIFKDLNL-KFAYQKGIGSYDITEVSAFISAFIGYYHG 235  
QY 226 VVGDHRRHVEIAYHPTALSDVPRTTASATLNTDYGWEIGRPF 269  
DB 236 VIGNKFEKIPV-ITPVVLNDAPQTTASATLNTDYGWEIGRPF 278

RESULT 13  
US-10-059-964-48  
; Sequence 48, Application US/10059964  
; Publication No. US20020120115A1  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko  
; APPLICANT: Ohai, No. US20020120115A1  
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
; TITLE OF INVENTION: Chafteensis  
; FILE REFERENCE: 22727/04021  
; CURRENT APPLICATION NUMBER: US/10/059,964  
; CURRENT FILING DATE: 2002-01-28  
; EARLIER APPLICATION NUMBER: 09/314,701  
; EARLIER FILING DATE: 1999-05-19  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
US-10-059-964-48

Query Match 35.5%; Score 509.5; DB 4; Length 280;  
Best Local Similarity 40.5%; Pred. No. 4.1e-42;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
QY 1 MNYKRFVVGVTLSFVFFLSDGAFSD---ANPSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLMSILPYQSPADPVGSRNDNKEGFYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIFALGLDKSEINTHNSNFTN-----SYDPTVASSFAGSGIIGYVNDFR 108  
DB 61 PINGTNSLTKKVFLGKKD-GDITKDDFTRVAPGIDFQNNLIS---GFSSGIGYSMDGPR 116  
QY 109 VEPGSGYENFEPEP--QWYPENSQSYKFFALSARNATNSDNKFIVLNNGVVDKSLNVNC 166  
DB 117 IELEAAVQOFPKNTDNDTNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLMVNTC 176



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 13:52:57 ; Search time 90 Seconds  
(without alignments)  
1323.018 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNVRFVGVTLSTFVFLS.....ASATLNTDYFGWEIGRPFAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_21.\*
- 1: Geneseq1980s.\*
  - 2: Geneseq1990s.\*
  - 3: Geneseq2000s.\*
  - 4: Geneseq2001s.\*
  - 5: Geneseq2002s.\*
  - 6: Geneseq2003as.\*
  - 7: Geneseq2003bs.\*
  - 8: Geneseq2004s.\*
  - 9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	271	5 AAU96118	Aau96118 Ehrlichia
2	1194	83.3	271	5 AAU73419	Aau73419 Ehrlichia
3	509.5	35.5	280	5 AAU96116	Aau96116 Ehrlichia
4	509.5	35.5	280	5 ABG77958	Abg77958 Ehrlichia
5	509.5	35.5	280	6 ADA09781	Ada09781 E. canis
6	509.5	35.5	280	9 ADW04274	Adw04274 Ehrlichia
7	507	35.4	283	5 AAU06944	Aay06944 E. chaffe
8	507	35.4	283	5 AAU96106	Aau96106 Ehrlichia
9	507	35.4	283	5 AAU73413	Aau73413 Ehrlichia
10	507	35.4	283	5 ABG77936	Abg77936 Ehrlichia
11	507	35.4	283	6 ADA09737	Ada09737 E. chaffe
12	507	35.4	283	9 ADW04230	Adw04230 Ehrlichia
13	494	34.4	285	2 AAY06957	Aay06957 E. chaffe
14	494	34.4	285	5 AAU73408	Aau73408 Ehrlichia
15	494	34.4	285	5 ABG77949	Abg77949 Ehrlichia
16	492	34.3	285	6 ADA09763	Ada09763 E. chaffe
17	492	34.3	285	9 ADW04256	Adw04256 Ehrlichia
18	489	34.1	279	2 AAY06954	Aay06954 E. chaffe
19	489	34.1	279	5 ABG77946	Abg77946 Ehrlichia
20	489	34.1	279	6 ADA09757	Ada09757 E. chaffe
21	489	34.1	279	9 ADW04250	Adw04250 Ehrlichia
22	482	33.6	279	5 AAU73405	Aau73405 Ehrlichia
23	463	32.3	283	2 AAY06955	Aay06955 E. chaffe
24	463	32.3	283	5 AAU73406	Aau73406 Ehrlichia

ALIGNMENTS

RESULT 1  
AAU96118  
ID AAU96118 standard; protein; 271 AA.

XX AAU96118;

AC AC

XX 02-JUL-2002 (first entry)

XX Ehrlichia canis p28-9.

XX Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.

XX Ehrlichia canis.

XX Ehrlichia canis.

XX WO200222782-A2.

XX 21-MAR-2002.

XX 12-SEP-2001; 2001WO-US028759.

XX 12-SEP-2000; 2000US-00660587.

XX (RERE-) RES DEV FOUND.

XX Walker DH, Yu X, McBride JW;

XX WPI; 2002-351882/38.

XX N-PSDB; ABK68878.

XX New recombinant homologous 28 kilodalton immunodominant protein from Ehrlichia canis, useful for treating Ehrlichia canis infections.

XX Claim 16; Fig 16; 106pp; English.

XX The invention relates to a recombinant homologous 28 kDa immunodominant protein, p28, (I), of Ehrlichia canis. (II), a 28-kDa antigen preferably dispersed in a pharmaceutically acceptable carrier, is useful for inhibiting E. canis infection in a subject. (I) is useful in the development of vaccines and serodiagnosis that are particularly effective for disease prevention and serodiagnosis. AAU96100-AAU96118 represent the 28-kDa antigen amino acid sequences of the invention

XX Sequence 271 AA;

Query Match 100.0%; Score 1434; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No. 9.3e-145; Indels 0; Gaps 0;  
Matches 271; Conservative 0; Mismatches 0

QY 1 MNYKRFVVGVTLTSTFFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
DB 1 MNYKRFVVGVTLTSTFFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSPAGFSGIIGYYVNDFRVEFGSYENFEP 120  
DB 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSPAGFSGIIGYYVNDFRVEFGSYENFEP 120  
QY 121 ERQWYFENSQSYKFFALSRNATSDNKFIVLENNGVWDKSLNNVCYDIASGSIPLAPYM 180  
DB 121 ERQWYFENSQSYKFFALSRNATSDNKFIVLENNGVWDKSLNNVCYDIASGSIPLAPYM 180  
QY 181 CAGVGADYIKFLGISLPKFSYQVKGYNPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPKFSYQVKGYNPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271  
DB 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

RESULT 2

AAU73419  
ID AAU73419 standard; protein; 271 AA.  
XX  
AC AAU73419;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Ehrlichia chaffeensis outer membrane protein P28-20.  
XX  
KW Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.  
XX  
OS Ehrlichia chaffeensis.  
XX  
PN WO200183699-A2.  
XX  
PD 08-NOV-2001.  
XX  
PF 01-MAY-2001; 2001WO-US013997.  
XX  
PR 01-MAY-2000; 2000US-0201035P.  
XX  
PA (RERE-) RES DEV FOUND.  
XX  
PI Walker DH, Yu X;  
XX  
PS WPI; 2002-066527/09.  
XX  
PT Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28  
XX  
PS useful as a vaccine against Ehrlichia chaffeensis.  
XX  
Claim 10; Fig 2; 97pp; English.

XX The invention relates to isolated and purified 28-kDa outer membrane  
CC proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are  
CC encoded by a 28kDa outer membrane protein multigene family. P28 proteins  
CC are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful  
CC for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia  
CC chaffeensis P28 outer membrane proteins of the invention  
XX  
SQ Sequence 271 AA;

Query Match 83.3%; Score 1194; DB 5; Length 271;  
Best Local Similarity 78.6%; Pred. No. 5.2e-119;  
Matches 213; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

QY 1 MNYKRFVVGVTLTSTFFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSABETIPG 60  
DB 1 MNYKRFVVGVALAILSLFLPDNSFSDANVPEGRGFVGVQYKVGIPNFSNFSABETILPG 60  
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSPAGFSGIIGYYVNDFRVEFGSYENFEP 120

DB 61 LTKSIFALGLDKSSISDHAGFTQAVNPTVASNFAGGVIIGYYVNDFRVEFGAYENFEP 120  
QY 121 ERQWYFENSQSYKFFALSRNATSDNKFIVLENNGVWDKSLNNVCYDIASGSIPLAPYM 180  
DB 121 ERQWYFEGGESHKFFALSRSTVQDNKFIVLENDGVIDKSLNNVCYDIAGSIPAPYM 180  
QY 181 CAGVGADYIKFLGISLPKFSYQVKGYNPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPKFSYQVKGYNPVSVNVMFLFGGYYHKVIGNRYERVEIAYHP 240  
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271  
DB 241 ATLTVNPKTTSASATLTDYFGWEVGMRTFL 271

RESULT 3

AAU96116  
ID AAU96116 standard; protein; 280 AA.  
XX  
AC AAU96116;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Ehrlichia canis p28-2.  
XX  
KW Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.  
XX  
OS Ehrlichia canis.  
XX  
PN WO200222782-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 12-SEP-2001; 2001WO-US028759.  
XX  
PR 12-SEP-2000; 2000US-00660587.  
XX  
PA (RERE-) RES DEV FOUND.  
XX  
PI Walker DH, Yu X, McBride JW;  
XX  
DR WPI; 2002-351882/38.  
XX  
DR N-PSDB; ABK68876.  
XX  
PT New recombinant homologous 28 kilodalton immunodominant protein from  
XX Ehrlichia canis, useful for treating Ehrlichia canis infections.  
XX  
PS Claim 16; Fig 14; 106pp; English.

XX The invention relates to a recombinant homologous 28 kDa immunodominant  
CC protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably  
CC dispersed in a pharmaceutically acceptable carrier, is useful for  
CC inhibiting E. canis infection in a subject. (I) is useful in the  
CC development of vaccines and serodiagnostics that are particularly  
CC effective for disease prevention and serodiagnosis. AAU96100-AAU96118  
CC represent the 28-kDa antigen amino acid sequences of the invention  
XX  
SQ Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 5; Length 280;  
Best Local Similarity 40.5%; Pred. No. 1.4e-45;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

QY 1 MNYKRFVVGVTLTSTFFVFLSDGAFSD---ANFSEGRGLYIGSQYKVGIPNFSNFSABET 57  
DB 1 MNYKILVRSALISLISILPYQSPADPVGSRDNDNKEGYISAKYNPSISHRKFSABET 60  
QY 58 -IPG---ITKKIFALGLDKSEINTHSNFTR-----SYDPTVASSPAGFSGIIGYYVNDFR 108  
DB 61 PINGNSLTAKVFGUKKO-GDITKKDDFTRVAPGIDFQNNLIS---GFSGSIQYSDMGPR 116  
QY 109 VEFEGSYENFEPER--QWYPENSQSYKFFALSRNATSDNKFIVLENNGVWDKSLNNVC 166

```

Db      117 IELEAAAYQQFNPKNNTDNDNGEYKHFALSRKDMEDQQYVVLKNDGITFSLMWNTC 176
Qy      167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPLNVNTMLFGGGYYHK 225
Db      177 YDITAGVSVFPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPIITPEVSFAFIGGYHG 235
Qy      226 VVGDHREVEIAHPTALSDVPRTTSSATLNTDYGWEIGFRF 269
Db      236 VIGNKFEKIPV-ITPVVLNDAPQTTASVTLDVGYFGGEIGMRF 278

```

## RESULT 4

```

ABG77958
ID      ABG77958 standard; protein; 280 AA.
XX
AC      ABG77958;
XX
DT      15-NOV-2002 (first entry)
XX
DE      Ehrlichia canis outer membrane protein (P30F) #9.
XX
KW      Outer membrane protein; OMP; P30F; ehrlichiosis; infection.
XX
OS      Ehrlichia canis.
XX
FN      US2002120115-A1.
XX
PD      29-AUG-2002.
XX
PF      28-JAN-2002; 2002US-00059964.
XX
PR      19-MAY-1999; 99US-00314701.
XX
PA      (RIKI/) RIKIHISA Y.
PA      (OHAS/) OHASHI N.
XX
PI      Rikihisa Y, Ohashi N;
XX
WPI; 2002-618954/66.
DR      N-PSDB; ABS63299.
XX

```

Isolated polynucleotide encoding an outer membrane protein of E.canis or E.chaffeensis used in the diagnosis of infection.

Claim 10; Fig 30B; 49pp; English.

The invention relates to an isolated polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used in the diagnosis of infection. An infection such as human ehrlichiosis or canine ehrlichiosis can be diagnosed by providing a serum sample from the patient, providing a polypeptide or mixture of polypeptides, contacting the sample with the polypeptide and assaying for the formation of a complex between antibodies in the serum sample and the polypeptide, where formation of a complex is indicative of infection with E. chaffeensis. This sequence represents an Ehrlichia outer membrane protein of the invention

Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 5; Length 280;  
Best Local Similarity 40.5%; Pred. No. 1.4e-45;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

```

Qy      1 MNYKRFVVGVTLSFTVFFLSDGAFSD---ANFSEGRGLYIGSQYKVGIPNFSNFASET 57
Db      1 MNYKKILVRSALISLSILPYQSFADPVGSRNDNKGFIYSAKNPSISHFRKFSABET 60
Qy      58 -IPG---ITKKIFALGLDKSEINTHSNFR-----SYDPTVASSPAGFSGIIGYVNDPR 108
Db      61 PINGTNSITKKVFGGLKGD-GDITKDDDFRVPAGDIFQNNLIS---GFGSGSIGYMDGPR 116
Qy      109 VEFEGSYENFEFER--QWYPENSQSYKFPALSRNATNSDNKPIVLNNGVVDKSLNVNC 166

```

```

Db      117 IELEAAAYQQFNPKNNTDNDNGEYKHFALSRKDMEDQQYVVLKNDGITFSLMWNTC 176
Qy      167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPLNVNTMLFGGGYYHK 225
Db      177 YDITAGVSVFPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPIITPEVSFAFIGGYHG 235
Qy      226 VVGDHREVEIAHPTALSDVPRTTSSATLNTDYGWEIGFRF 269
Db      236 VIGNKFEKIPV-ITPVVLNDAPQTTASVTLDVGYFGGEIGMRF 278

```

## RESULT 5

```

ADA09781
ID      ADA09781 standard; protein; 280 AA.
XX
AC      ADA09781;
XX
DT      06-NOV-2003 (first entry)
XX
DE      E. canis outer membrane protein P30-10.
XX
KW      outer membrane protein; circulating leukocyte; monocytic ehrlichiosis; Rocky Mountain spotted fever; canine ehrlichiosis; antigen.
XX
OS      Ehrlichia canis.
XX
FN      US65444517-B1.
XX
PD      08-APR-2003.
XX
PF      19-MAY-1999; 99US-00314701.
XX
PR      18-SEP-1998; 98US-0100843P.
XX
PA      (OHIS ) UNIV OHIO STATE RES FOUND.
XX
PI      Rikihisa Y, Ohashi N;
XX
WPI; 2003-553952/52.
DR      N-PSDB; ADA09780.
XX

```

New isolated polynucleotide encoding outer membrane protein P30 of Ehrlichia canis or its variant or fragment, useful for producing Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for diagnosing and treating ehrlichiosis.

Disclosure; Fig 30; 105pp; English.

The invention relates to an isolated polynucleotide encoding a variant of the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer membrane protein of E. canis, or an antigenic fragment of the E. canis P30 protein, or comprising a sequence which is the complement of nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E. canis and E. chaffeensis outer membrane proteins and their encoding nucleic acids. The polynucleotides are useful for producing E. canis or E. chaffeensis outer membrane protein, for designing hybridisation probes for isolating and identifying cDNA and genomic clones encoding the OMP or its allelic forms, for designing primers for PCR. The polypeptides encoded by the polynucleotide is useful for diagnosing human ehrlichiosis (Rocky Mountain spotted fever) or canine ehrlichiosis. The present sequence represents an E. canis outer membrane protein.

Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 6; Length 280;  
Best Local Similarity 40.5%; Pred. No. 1.4e-45;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

```

Qy      1 MNYKRFVVGVTLSFTVFFLSDGAFSD---ANFSEGRGLYIGSQYKVGIPNFSNFASET 57
Db      1 MNYKKILVRSALISLSILPYQSFADPVGSRNDNKGFIYSAKNPSISHFRKFSABET 60

```

QY 58 -IPG---ITKKIFALGLDKSEINTHSNFR-----SYDPTYASSPAGFGSGIIGYVNDPR 108  
 Db 61 PINGTNSLTKKVFGLKKD-GDITKKDDFTFVAPGIDFQNNLIS--GFSGSIGYSMDGPR 116  
 QY 109 VEPFGSYENFEPER--QWYPENSQSKYKFPALSRNATNSDNKFTVLNNGVVDKSLNVNVC 166  
 Db 117 IELEAAAYQQFNPKNTDNDNGEYKHFALSRKDAMEQQYVVLKNDGITEWMLMVNTC 176  
 QY 167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPSKFSYQVKGVNYPLNVNMTLPGGGYHYK 225  
 Db 177 YDITAGVSVFPYACAGIGADLITIFKQNLNLFKAYQGGKIGISYPTTPEVSAPFIGGYHYG 235  
 QY 226 VVGDHREHVEIAYHPTALSVDPRTTTSASATLNTDYFGWEIGPRF 269  
 Db 236 VIGNKFEKIPV-ITPVVLNDAPQTTASVTLDVGYFGGEIGMRF 278

RESULT 6  
 ID ADW04274  
 XX ADW04274 standard; protein; 280 AA.  
 AC ADW04274;  
 XX  
 DT 24-MAR-2005 (first entry)  
 DE Ehrlichia canis outer membrane protein (P30F), P30-10.  
 XX  
 KW DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;  
 KW infection; vaccine.  
 XX  
 OS Ehrlichia canis.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide I. .25  
 FT /label= Signal\_peptide  
 FT Protein 26..280  
 FT /note= "Ehrlichia canis mature OMP protein"  
 FT  
 XX US2004265334-A1.  
 XX  
 PD 30-DEC-2004.  
 XX  
 XX 29-JUL-2004; 2004US-00901774.  
 XX  
 PR 18-SEP-1998; 98US-0100843P.  
 PR 19-MAY-1999; 99US-00314701.  
 PR 28-JAN-2002; 2002US-00059964.  
 XX  
 XX (RIKI/) RIKIHISA Y.  
 XX (OHAS/) OHASHI N.  
 XX  
 PI Rikihisa Y, Ohashi N;  
 XX  
 DR WPI; 2005-064871/07.  
 DR N-PSDB; ADW04273.  
 XX  
 XX New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia  
 PT canis or Ehrlichia chaffeensis, useful in preparing a composition for  
 PT diagnosing or preventing E. canis or E. chaffeensis infection.  
 XX  
 XX Claim 10; SEQ ID NO 46; 122pp; English.  
 XX  
 XX The invention relates to nucleic acid sequences encoding outer membrane  
 CC proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and  
 CC Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is  
 CC useful in preparing a composition for diagnosing, treating or preventing  
 CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present  
 CC sequence is the Ehrlichia canis P30F protein.  
 XX  
 SQ Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 9; Length 280;  
 Best Local Similarity 40.5%; Pred. No. 1.4e-45;

Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
 QY 1 MNYKFPVGVVTLSTVFFFLSDGAFSD--ANPSEGRGLYIGSQYKVGIPNFSNFAEET 57  
 Db 1 MNYKKILVRSALISLMSILPYQSPADPVGSRDNDNKGIFYSAKNPISISHFRKFSAEET 60  
 QY 58 -IPG---ITKKIFALGLDKSEINTHSNFR-----SYDPTYASSPAGFGSGIIGYVNDPR 108  
 Db 61 PINGTNSLTKKVFGLKKD-GDITKKDDFTFVAPGIDFQNNLIS--GFSGSIGYSMDGPR 116  
 QY 109 VEPFGSYENFEPER--QWYPENSQSKYKFPALSRNATNSDNKFTVLNNGVVDKSLNVNVC 166  
 Db 117 IELEAAAYQQFNPKNTDNDNGEYKHFALSRKDAMEQQYVVLKNDGITEWMLMVNTC 176  
 QY 167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPSKFSYQVKGVNYPLNVNMTLPGGGYHYK 225  
 Db 177 YDITAGVSVFPYACAGIGADLITIFKQNLNLFKAYQGGKIGISYPTTPEVSAPFIGGYHYG 235  
 QY 226 VVGDHREHVEIAYHPTALSVDPRTTTSASATLNTDYFGWEIGPRF 269  
 Db 236 VIGNKFEKIPV-ITPVVLNDAPQTTASVTLDVGYFGGEIGMRF 278

RESULT 7  
 ID AAY06944  
 XX AAY06944 standard; protein; 283 AA.  
 AC AAY06944;  
 XX

DT 27-AUG-2003 (revised)  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE E. chaffeensis OMP-1B protein.  
 XX  
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;  
 KW detection; dog.  
 XX  
 OS Ehrlichia chaffeensis.  
 XX  
 XX WO9913720-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 18-SEP-1998; 98WO-US019600.  
 XX  
 PR 19-SEP-1997; 97US-0059353P.  
 XX  
 XX (OHIS ) UNIV OHIO STATE.  
 XX  
 XX Rikihisa Y, Ohashi N;  
 XX  
 DR WPI; 1999-254290/21.  
 DR N-PSDB; AAX34744.  
 XX  
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia  
 PT canis.  
 XX  
 XX Claim 12; Fig 4B; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from  
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part  
 CC of the OMP family and consist of proteins OMP-1, -1(B to 2) shown in  
 CC AAY06943-958. The E. canis proteins form part of the P30 family and  
 CC consist of proteins shown in AAY06959-970. The proteins and genes are  
 CC used to detect E. chaffeensis in patients and E. canis in dogs. (Updated  
 CC on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 283 AA;

Query Match 35.4%; Score 507; DB 2; Length 283;  
 Best Local Similarity 40.3%; Pred. No. 2.7e-45;  
 Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;



```

Db      119  RIELEAAAYQKFDKXN---PDNDNTNSGDYKYFGLSREDAIADKKVVLKNEGITPMSLM 175
QY      163  VNVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGWYPLNVNTMLFGGG 221
Db      176  VNTCYDITABGVFFIPIYACAGVADLINVFKDFNL-KFSYQKGIGISYPTITPEVSAPFIG 234
QY      222  YHKVVGDRHERVEIAVHPALSDVPRTTSSASATLNTDYGWGEIGPRF 269
Db      235  YHGVIGNNFNKIPV-ITPVVLEGAPOTTLSALVTIDTGYFGGEGVGRF 281

```

## RESULT 10

ABG77936  
ID ABG77936 standard; protein; 283 AA.

XX  
AC ABG77936;

XX  
DT 15-NOV-2002 (first entry)

XX  
DE Ehrlichia chaffeensis outer membrane protein (OMP) #2.

XX  
KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.

XX  
OS Ehrlichia chaffeensis.

XX  
PN US2002120115-A1.

XX  
PD 29-AUG-2002.

XX  
PF 28-JAN-2002; 2002US-00059964.

XX  
PR 19-MAY-1999; 99US-00314701.

XX  
PA (RIKI/) RIKIHISA Y.

XX  
PA (OHAS/) OHASHI N.

XX  
PI Rikihisa Y, Ohashi N;

XX  
DR WPI; 2002-618954/66.

XX  
DR N-PSDB; ABS63277.

XX  
PT Isolated polynucleotide encoding an outer membrane protein of E.canis or  
PT E.chaffeensis used in the diagnosis of infection.

XX  
PS Disclosure; Fig 4B; 49pp; English.

XX  
CC The invention relates to an isolated polynucleotide encoding an outer  
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
CC in the diagnosis of infection. An infection such as human ehrlichiosis or  
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
CC patient, providing a polypeptide or mixture of polypeptides, contacting  
CC the sample with the polypeptide and assaying for the formation of a  
CC complex between antibodies in the serum sample and the polypeptide, where  
CC formation of a complex is indicative of infection with E. chaffeensis.  
CC This sequence represents an Ehrlichia outer membrane protein of the  
CC invention

XX  
SQ Sequence 283 AA;

Query Match 35.4%; Score 507; DB 5; Length 283;  
Best Local Similarity 40.3%; Pred. No. 2.7e-45;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

```

QY      1  MNYKRFVVGVTLSFTVFVFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54
Db      1  MNYKKIFVSSALISLSILFYQSFDAPVTSNDTGINDSREGFYISVKYNPSISHFRKFA 60

```

```

QY      55  EET-IPG---ITKKIFALGLDKSEINTHSNFTSYDPTYA---SSPAGFGSIIGYVYVDF 107

```

```

Db      61  EEAPINGNTSITKKVFLGKKD-GDIAQSANFNT-DPALEFQNLISGFSGSLGYANDGP 118

```

```

QY      108  RVEFEGSYENFEPERQWYPENSQS-----YKFPALSRNATNSDNKFTVLNNGVVDKSLN 162

```

```

Db      119  RIELEAAAYQKFDKXN---PDNDNTNSGDYKYFGLSREDAIADKKVVLKNEGITPMSLM 175
QY      163  VNVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGWYPLNVNTMLFGGG 221
Db      176  VNTCYDITABGVFFIPIYACAGVADLINVFKDFNL-KFSYQKGIGISYPTITPEVSAPFIG 234
QY      222  YHKVVGDRHERVEIAVHPALSDVPRTTSSASATLNTDYGWGEIGPRF 269
Db      235  YHGVIGNNFNKIPV-ITPVVLEGAPOTTLSALVTIDTGYFGGEGVGRF 281

```

## RESULT 11

ADA09737

ID ADA09737 standard; protein; 283 AA.

XX  
AC ADA09737;

XX  
DT 06-NOV-2003 (first entry)

XX  
DE E. chaffeensis outer membrane protein OMP-1B.

XX  
KW outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;  
KW Rocky Mountain spotted fever; canine ehrlichiosis; antigen.

XX  
OS Ehrlichia chaffeensis.

XX  
PN US6544517-B1.

XX  
PD 08-APR-2003.

XX  
PF 19-MAY-1999; 99US-00314701.

XX  
PR 18-SEP-1998; 98US-0100843P.

XX  
PA (OHIS ) UNIV OHIO STATE RES FOUND.

XX  
PI Rikihisa Y, Ohashi N;

XX  
DR WPI; 2003-553952/52.

XX  
DR N-PSDB; ADA09736.

XX  
PT New isolated polynucleotide encoding outer membrane protein P30 of  
PT Ehrlichia canis or its variant or fragment, useful for producing  
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for  
PT diagnosing and treating ehrlichiosis.

XX  
PS Disclosure; Fig 4; 105pp; English.

XX  
CC The invention relates to an isolated polynucleotide encoding a variant of  
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer  
CC membrane protein of E. canis , or an antigenic fragment of the E. canis  
CC P30 protein, or comprising a sequence which is the complement of  
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.  
CC canis and E. chaffeensis outer membrane proteins and their encoding  
CC nucleic acids. The polynucleotides are useful for producing E. canis or  
CC E. chaffeensis outer membrane protein, for designing hybridisation probes  
CC for isolating and identifying cDNA and genomic clones encoding the OMP)  
CC or its allelic forms, for designing primers for PCR. The polypeptides  
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis  
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present  
CC sequence represents an E. chaffeensis outer membrane protein.

XX  
SQ Sequence 283 AA;

Query Match 35.4%; Score 507; DB 6; Length 283;  
Best Local Similarity 40.3%; Pred. No. 2.7e-45;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

```

QY      1  MNYKRFVVGVTLSFTVFVFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54

```

```

Db      1  MNYKKIFVSSALISLSILFYQSFDAPVTSNDTGINDSREGFYISVKYNPSISHFRKFA 60

```





CC on 27-AUG-2003 to correct OS field.)

XX Sequence 285 AA;

Query Match 34.4%; Score 494; DB 2; Length 285;  
Best Local Similarity 38.4%; Pred. No. 6.8e-44;  
Matches 111; Conservative 57; Mismatches 95; Indels 26; Gaps 10;

QY 1 MNYKR--FVVGVTLSFVFLSDGAFSD--ANFSEGRGLYIGSOYKVGIPNFSNESAEE 56  
DB 1 MNNRKSFFIIGASLLASLLTSEASTGNVSNHTFKPRLYISGQVRPGVSHFSKFSVKE 60  
QY 57 T-----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASSPAGFSGIIGY-YVNDFRV 109  
DB 61 TNYNTTQLVGLKKDISVIG--NSNITTNTNFNFPYIAEFQDNALISFGAIGVLYSENFR 118  
QY 110 EFGSYENFEPERQWYVENS--QSYKFFALSR-----NATNSDN--KFIVLENNGVVDKS 160  
DB 119 EVEASYEEFVDVKN---PEGSATDAYRYFALARAMDGTNKSPPDTRKFTVMRNDGLSIS 175  
QY 161 LNVNVCYDIASGISPLAPYMCAGVADYIKFLGSLPKFSYQVKGVNYPLNVNTMLFGG 220  
DB 176 VMINGCYNFTLDDIPVVPVVCAGIGDFIEFFNDLHVKAHQKVGISYSISPEVSLFLN 235  
QY 221 GYHKVGVDRHERVEIAYHPTALSDVPRITTSASATLNTDYFGWEIGFRF 269  
DB 236 GYHKVGTGNRFKNLHV-QHVS DLSAPKFTSAVATLNVGVFGGEIGVRF 283

RESULT 14

AAU73408  
ID AAU73408 standard; protein; 285 AA.

AC AAU73408;

DT 12-MAR-2002 (first entry)

DE Ehrlichia chaffeensis outer membrane protein P28-9.

KW Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.

OS Ehrlichia chaffeensis.

PN WO200183699-A2.

PD 08-NOV-2001.

PF 01-MAY-2001; 2001WO-US013997.

PR 01-MAY-2000; 2000US-0201035P.

PA (RERE-) RES DEV FOUND.

PI Walker DH, Yu X;

DR WPI; 2002-066527/09.

PT Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28  
useful as a vaccine against Ehrlichia chaffeensis.

PS Claim 10; Fig 2; 97pp; English.

CC The invention relates to isolated and purified 28-kDa outer membrane  
proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are  
encoded by a 28KDa outer membrane protein multigene family. P28 proteins  
are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful  
for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia  
chaffeensis P28 outer membrane proteins of the invention

XX Sequence 285 AA;

Query Match 34.4%; Score 494; DB 5; Length 285;  
Best Local Similarity 38.4%; Pred. No. 6.8e-44;

Matches 111; Conservative 57; Mismatches 95; Indels 26; Gaps 10;

QY 1 MNYKR--FVVGVTLSFVFLSDGAFSD--ANFSEGRGLYIGSOYKVGIPNFSNESAEE 56  
DB 1 MNNRKSFFIIGASLLASLLTSEASTGNVSNHTFKPRLYISGQVRPGVSHFSKFSVKE 60  
QY 57 T-----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASSPAGFSGIIGY-YVNDFRV 109  
DB 61 TNYNTTQLVGLKKDISVIG--NSNITTNTNFNFPYIAEFQDNALISFGAIGVLYSENFR 118  
QY 110 EFGSYENFEPERQWYVENS--QSYKFFALSR-----NATNSDN--KFIVLENNGVVDKS 160  
DB 119 EVEASYEEFVDVKN---PEGSATDAYRYFALARAMDGTNKSPPDTRKFTVMRNDGLSIS 175  
QY 161 LNVNVCYDIASGISPLAPYMCAGVADYIKFLGSLPKFSYQVKGVNYPLNVNTMLFGG 220  
DB 176 VMINGCYNFTLDDIPVVPVVCAGIGDFIEFFNDLHVKAHQKVGISYSISPEVSLFLN 235  
QY 221 GYHKVGVDRHERVEIAYHPTALSDVPRITTSASATLNTDYFGWEIGFRF 269  
DB 236 GYHKVGTGNRFKNLHV-QHVS DLSAPKFTSAVATLNVGVFGGEIGVRF 283

RESULT 15

ABG77949  
ID ABG77949 standard; protein; 285 AA.

XX AC ABG77949;

XX DT 15-NOV-2002 (first entry)

XX DE Ehrlichia chaffeensis outer membrane protein (OMP) #15.

KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.

OS Ehrlichia chaffeensis.

PN US2002120115-A1.

PD 29-AUG-2002.

PF 28-JAN-2002; 2002US-00059964.

PR 19-MAY-1999; 99US-00314701.

PA (RIKI/) RIKIHISA Y.

PA (OHAS/) OHASHI N.

PI Rikihisa Y, Ohashi N;

DR WPI; 2002-618954/66.

DR N-PSDB; ABS63290.

PT Isolated polynucleotide encoding an outer membrane protein of E.canis or  
E.chaffeensis used in the diagnosis of infection.

PS Claim 14; Fig 17B; 49pp; English.

CC The invention relates to an isolated polynucleotide encoding an outer  
membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
in the diagnosis of infection. An infection such as human ehrlichiosis or  
canine ehrlichiosis can be diagnosed by providing a serum sample from the  
patient, providing a polypeptide or mixture of polypeptides, contacting  
the sample with the polypeptide and assaying for the formation of a  
complex between antibodies in the serum sample and the polypeptide, where  
formation of a complex is indicative of infection with E. chaffeensis.  
This sequence represents an Ehrlichia outer membrane protein of the  
invention

XX Sequence 285 AA;

Query Match 34.4%; Score 494; DB 5; Length 285;  
Best Local Similarity 38.4%; Pred. No. 6.8e-44;

	Matches	111;	Conservative	57;	Mismatches	95;	Indels	26;	Gaps	10;
Qy	1	MNYKR--FVVGVTLTSTVFFLSDGAFSD--ANFSSEGRGLYIGSQYKVGIPNFSNPSAAEE	56							
Db	1	NNNRKSFFIIAGSLASLLLETTSETAASSTGNVSNHTYFKPRLYISQSYPGVSHSFKFSVKKE	60							
Qy	57	T-----IPGITTKIIFALGLDKSINTHSNPTRSYDPTYASSPAGFGSIGIIGY-VYNDPRV	109							
Db	61	TNYNTTOLVGLKKDIISVIG--NSNIITYTNFFPIAEFPQDNALSFSGAIGLYSENFR	118							
Qy	110	EFEGSYNPFPEROWYPENS--QYKPPALSR-----NATNSDN--KFTVLNNNGVVVDKS	160							
Db	119	EVEASYBEFDVKN---PEGSATDAYRYFALARAMDGTNKSPDDTRKFTVMRNDGLSISS	175							
Qy	161	LNVNVCYDIDIASGISPLAPYMCAGVGADYIKFLGILSKPKFSQVYKFVNYPPLNVNMTLFGG	220							
Db	176	VMINGCYNFLTLDIPVPVYPVCANGIGGGDFIEFFNDLHVKAHQGVGISYSISPEVSFLNL	235							
Qy	221	GYHKVGDGRHVERVEIAHYHPALTSDVRPTTSASATLNTDYFGWIGIRFP	269							
Db	236	GYHKVTGNFKNLHV-QHVSDDLSDAPKFTSAVALNTXVGPFGSGIVRF	283							

Search completed: December 28, 2005, 15:36:01  
Job time : 92 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 10:46:11 ; Search time 197 Seconds  
(without alignments)  
7335.826 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaaaagattgt.....ttggatttagattgcgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCITUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813	100.0	813	3	US-09-660-587-45
2	813	100.0	813	3	US-09-811-007A-45
3	162.2	20.0	840	3	US-09-314-701-23
4	162.2	20.0	840	3	US-10-314-639-23
5	162.2	20.0	840	3	US-10-059-964A-23
6	150.8	18.5	852	3	US-09-314-701-25
7	150.8	18.5	852	3	US-10-314-639-25
8	150.8	18.5	852	3	US-10-059-964A-25
9	147.2	18.1	840	3	US-09-660-587-41
10	147.2	18.1	840	3	US-09-811-007A-41
11	147.2	18.1	843	3	US-09-314-701-47
12	147.2	18.1	843	3	US-10-314-639-47
13	147.2	18.1	843	3	US-10-059-964A-47
14	146.6	18.0	726	3	US-09-314-701-61
15	146.6	18.0	726	3	US-10-314-639-61
16	146.6	18.0	726	3	US-10-059-964A-61
17	146.4	18.0	849	3	US-09-648-520B-48
18	135	16.6	828	3	US-09-314-701-27
19	135	16.6	828	3	US-10-314-639-27
20	135	16.6	828	3	US-10-059-964A-27
21	133.8	16.5	840	3	US-09-314-701-59
22	133.8	16.5	840	3	US-10-314-639-59
23	133.8	16.5	840	3	US-10-059-964A-59
24	129.6	15.9	843	3	US-09-314-701-5

25	129.6	15.9	843	3	US-10-314-639-5	Sequence 5, Appli
26	129.6	15.9	843	3	US-10-059-964A-5	Sequence 5, Appli
27	126.4	15.5	837	3	US-08-953-326-9	Sequence 9, Appli
28	126.4	15.5	837	3	US-09-314-701-9	Sequence 9, Appli
29	126.4	15.5	837	3	US-09-553-662-9	Sequence 9, Appli
30	126.4	15.5	837	3	US-10-062-994-9	Sequence 9, Appli
31	126.4	15.5	837	3	US-10-314-639-9	Sequence 9, Appli
32	126.4	15.5	837	3	US-10-059-964A-9	Sequence 9, Appli
33	122.4	15.1	867	3	US-09-314-701-31	Sequence 31, Appl
34	122.4	15.1	867	3	US-10-314-639-31	Sequence 31, Appl
35	122.4	15.1	867	3	US-10-059-964A-31	Sequence 10, Appl
36	121.8	15.0	843	3	US-08-953-326-10	Sequence 10, Appl
37	121.8	15.0	843	3	US-09-314-701-11	Sequence 11, Appl
38	121.8	15.0	843	3	US-09-553-662-10	Sequence 10, Appl
39	121.8	15.0	843	3	US-10-062-994-10	Sequence 10, Appl
40	121.8	15.0	843	3	US-10-314-639-11	Sequence 11, Appl
41	121.8	15.0	843	3	US-10-059-964A-11	Sequence 11, Appl
42	120.8	14.9	897	3	US-09-314-701-51	Sequence 51, Appl
43	120.8	14.9	897	3	US-10-314-639-51	Sequence 51, Appl
44	120.8	14.9	897	3	US-10-059-964A-51	Sequence 51, Appl
45	118.2	14.5	852	3	US-09-314-701-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-660-587-45  
; Patent No. 6392023  
; Sequence 45, Application US/09660587  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: YU, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; TITLE OF INVENTION: Genes Of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 45  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9  
US-09-660-587-45

Query Match	100.0%	Score 813	DB 3	Length 813	Matches 813	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGAATACAAAAGATTGTTGAGGTGTTACGCTGAGTACATTTGTTTCTTCTATCT	60						
Db	1	ATGAATACAAAAGATTGTTGAGGTGTTACGCTGAGTACATTTGTTTCTTCTATCT	60						
Qy	61	GATGGTCTTTTCTGATGCAATTTTCTGAAGGGAGGAGGACCTTTATATAGTAGT	120						
Db	61	GATGGTCTTTTCTGATGCAATTTTCTGAAGGGAGGAGGACCTTTATATAGTAGT	120						
Qy	121	CAGTATAAAGTTGGTATTTCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT	180						
Db	121	CAGTATAAAGTTGGTATTTCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT	180						
Qy	181	ATTACAAAAAGATTGTTGAGGTGTTGATAGTCTGAGATAAATACTACAGCAAT	240						
Db	181	ATTACAAAAAGATTGTTGAGGTGTTGATAGTCTGAGATAAATACTACAGCAAT	240						
Qy	241	TTTACAGCATATGACCTTACTTATGCAAGCAGTTTTCAGGGTTTAGTGGTATCAT	300						
Db	241	TTTACAGCATATGACCTTACTTATGCAAGCAGTTTTCAGGGTTTAGTGGTATCAT	300						

QY 301 GGATATTATGTTAAAGCTTTAGGCTAGAAATTTTGAAGGTTCTTATGAGAAATTTTGAACCT 360  
 DB |||||  
 DB 301 GGATATTATGTTAAAGCTTTAGGCTAGAAATTTTGAAGGTTCTTATGAGAAATTTTGAACCT 360  
 QY 361 GAAAGCAATGGTACCTCGAGATAGCCAAAGCTACAAAATTTTGTCTCGAAAT 420  
 DB |||||  
 DB 361 GAAAGCAATGGTACCTCGAGATAGCCAAAGCTACAAAATTTTGTCTCGAAAT 420  
 QY 421 GCTACAAATAGTGATTAATAGTTTATAGTACTAGAGAAATACGGCGTTGTCACAAGTCT 480  
 DB |||||  
 DB 421 GCTACAAATAGTGATTAATAGTTTATAGTACTAGAGAAATACGGCGTTGTCACAAGTCT 480  
 QY 481 CTTAAATGTAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
 DB |||||  
 DB 481 CTTAAATGTAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
 QY 541 TGTGCTGGTGTGGTGCAGATTAATATAAAGTTTATAGGTATATCATTTGCTAAGTTTTCT 600  
 DB |||||  
 DB 541 TGTGCTGGTGTGGTGCAGATTAATATAAAGTTTATAGGTATATCATTTGCTAAGTTTTCT 600  
 QY 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAAATGTTTAAATGTTTGGTGG 660  
 DB |||||  
 DB 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAAATGTTTAAATGTTTGGTGG 660  
 QY 661 GGTATTACCATAAAGTTTGTAGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCT 720  
 DB |||||  
 DB 661 GGTATTACCATAAAGTTTGTAGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCT 720  
 QY 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780  
 DB |||||  
 DB 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780  
 QY 781 TTTGGTTGGAGATTGGATTGGCGCTA 813  
 DB |||||  
 DB 781 TTTGGTTGGAGATTGGATTGGCGCTA 813

RESULT 2

US-09-811-007A-45  
 ; Sequence 45, Application US/09811007A  
 ; Patent No. 660269  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walker, David H.  
 ; APPLICANT: McBride, Jere W.  
 ; APPLICANT: Yu, Xue-Jie  
 ; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
 ; FILE REFERENCE: D6152CIP2  
 ; CURRENT APPLICATION NUMBER: US/09/811,007A  
 ; PRIOR FILING DATE: 2001-10-23  
 ; PRIOR FILING DATE: 2000-09-12  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SEQ ID NO 45  
 ; LENGTH: 813  
 ; TYPE: DNA  
 ; ORGANISM: Ehrlichia canis  
 ; FEATURE:  
 ; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9  
 US-09-811-007A-45

Query Match 100.0%; Score 813; DB 3; Length 813;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-197;  
 Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAATTACAAAGATTGTTAGGTGTACCTCGAGTACATTTGTTTTTCTTATCT 60  
 DB |||||  
 DB 1 ATGAATTACAAAGATTGTTAGGTGTACCTCGAGTACATTTGTTTTTCTTATCT 60  
 QY 61 GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATATAGTACT 120  
 DB |||||  
 DB 61 GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATATAGTACT 120

QY 121 CAGTATAAAGTTGGTATTTCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTTCTGCT 180  
 DB |||||  
 DB 121 CAGTATAAAGTTGGTATTTCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTTCTGCT 180  
 QY 181 ATTACAAAAAGATTTTTGGCTTCTGTATAGTCTGTGATTAAGTCTGAGATAAATACTCACAGCAAT 240  
 DB |||||  
 DB 181 ATTACAAAAAGATTTTTGGCTTCTGTATAGTCTGTGATTAAGTCTGAGATAAATACTCACAGCAAT 240  
 QY 241 TTTACACATCATATGACCCCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGGTATCATTT 300  
 DB |||||  
 DB 241 TTTACACATCATATGACCCCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGGTATCATTT 300  
 QY 301 GGATATTATGTTAAATGACTTTTAGGGTAGAAATTTGAAAGTTCTTATGAGAAATTTTGAACCT 360  
 DB |||||  
 DB 301 GGATATTATGTTAAATGACTTTTAGGGTAGAAATTTGAAAGTTCTTATGAGAAATTTTGAACCT 360  
 QY 361 GAAAGCAATGGTACCTCGAGATAGCCAAAGCTACAAAATTTTGTCTCGAAAT 420  
 DB |||||  
 DB 361 GAAAGCAATGGTACCTCGAGATAGCCAAAGCTACAAAATTTTGTCTCGAAAT 420  
 QY 421 GCTACAAATAGTGATTAATAGTTTATAGTACTAGAGAAATACGGCGTTGTCACAAGTCT 480  
 DB |||||  
 DB 421 GCTACAAATAGTGATTAATAGTTTATAGTACTAGAGAAATACGGCGTTGTCACAAGTCT 480  
 QY 481 CTTAAATGTAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
 DB |||||  
 DB 481 CTTAAATGTAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
 QY 541 TGTGCTGGTGTGGTGCAGATTAATATAAAGTTTATAGGTATATCATTTGCTAAGTTTTCT 600  
 DB |||||  
 DB 541 TGTGCTGGTGTGGTGCAGATTAATATAAAGTTTATAGGTATATCATTTGCTAAGTTTTCT 600  
 QY 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAAATGTTTAAATGTTTGGTGG 660  
 DB |||||  
 DB 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAAATGTTTAAATGTTTGGTGG 660  
 QY 661 GGTATTACCATAAAGTTTGTAGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCT 720  
 DB |||||  
 DB 661 GGTATTACCATAAAGTTTGTAGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCT 720  
 QY 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780  
 DB |||||  
 DB 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780  
 QY 781 TTTGGTTGGAGATTGGATTGGCGCTA 813  
 DB |||||  
 DB 781 TTTGGTTGGAGATTGGATTGGCGCTA 813

RESULT 3

US-09-314-701-23  
 ; Sequence 23, Application US/09314701  
 ; Patent No. 6544517  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rikihisa, Yasuko  
 ; APPLICANT: Ohasi, No. 6544517io  
 ; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
 ; FILE REFERENCE: 22727/04021  
 ; CURRENT APPLICATION NUMBER: US/09/314,701  
 ; CURRENT FILING DATE: 1999-05-19  
 ; NUMBER OF SEQ ID NOS: 66  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 23  
 ; LENGTH: 840  
 ; TYPE: DNA  
 ; ORGANISM: Ehrlichia chaffeensis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(840)  
 US-09-314-701-23

Query Match	20.0%	Score 162.2	DB 3	Length 840
Best Local Similarity	53.8%	Pred. No. 9.4e-32		
Matches	Conservative 0	Mismatches 353	Indels 18	Gaps 4
QY	22	GTAGGTGTTACGCTGAGTACATATTTGTTTTTTCTTATCTGATGGTGCTTTTTCTCATGCA	81	
DB	34	GTACTTGGCATCTCTATATCATCTTATCTATTTGAACTCCTTTTCAGCTATAAATCATAT	93	
QY	82	AATTTTTCTGAAGGAGGAGAGGACTTTATATATAGGTAGTCAGTATAAAGTTGGTATCCC	141	
DB	94	CATACAGAAATTAACACTAGTGGTATATATATACAGGCGAGTATAGACGAGGATATCC	153	
QY	142	RAATTTTTAGTAATTTTTTCAGCTGGAAGAACAACTCTCTGGTATTACAAAAAAGATTTTTGGC	201	
DB	154	CATTTTTGCAATTTCTCAGTAAAGAACTAAATGTTGATCAATACACTAGTAGGATAT	213	
QY	202	TTAGTCTTTGATGAAGTCTCT--GAGATAAATCTCACAGCAATTTTACAGATCATATGAC	258	
DB	214	AAAAAAGTGGCTCTCTATCGATCTTAAACACTATTTCAAACTTTCAAGTCCATATACT	273	
QY	259	CTACTTATGCAGCAGTTTTGCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGAC	318	
DB	274	GTTCATTTCAAGATAATGCTGCTAGTTCAGTGGAGCAATTTGGATATTCTTACCCCGAA	333	
QY	319	TTT---AGGCTAGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCTCGAAAGACAAATGGTAC	375	
DB	334	AGTCTAAGACTTTGAACCTTTGAAGGTTCTTACGAAAAATTTGATGCTCAAGATCCTAAAGAC	393	
QY	376	CTGAGAAATAGCAGCAAGCTACAAATTTTTTGTCTTGTCTCGAAATGCTACAAATAGTGTAT	435	
DB	394	TAICTCAGCAAAAGATGCTTTTTAGTGTTTTTTGTCTAGCAGTAATACGCTACTACTGTGT	453	
QY	436	AAT-----AAGTTTTATAGTACTAGAGAAATACGGCGTGTTCGCAAGTCTCTTAAT	486	
DB	454	CTGATGCTCAAAATATACAGTTATGAGAAATATGGCTTATCTGTTGGATCAATCATG	513	
QY	487	GTAAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCT	546	
DB	514	ATCAATGGTTGTATGATCTATCTTTTAAATAATTTAGTCGTATCACCCTTATATATGCA	573	
QY	547	GGTGTGGTGCAGATATATAAGTTTTTATAGGTATATCATTTGCTAAGTTTTCTTTATCAA	606	
DB	574	GGTATGGTGAAGATTTCATTTGAATTTTTTGTACTTTGCACATTAACCTTGTCTATCAA	633	
QY	607	GTTAAAGTTGGTGTCAACTACCTCTAAATGTTAAATACTATGTTGTTGGTGGGGTTAT	666	
DB	634	GGAAACTAGGTATTAGTTATTACTTCTTCCTAAGATTAATGTAATTTGCTGGTGGTAC	693	
QY	667	TACCAATAAGGTTGTAGGTGATAGGCATGAGAGAGTAGAAATAGCTTACCAATCCTACTGCA	726	
DB	694	TATCATAGAGTTATAGGGAATAAATTTAAAAATTTT---AAATGTTAAACCATGTTGTACA	750	
QY	727	TTATCTGAGGTTCTTGAAGAACTTTCAGCTTCTCTACTTTTAAATACTGATATTATTTGTT	786	
DB	751	CTTGATGAATTTCTTAAGCAACTTCTGCAGTAGCTACCTTAATGTTGCTTATTTTGGT	810	
QY	787	TGGAGATTTGGATTTTGAATTTGTC	809	
DB	811	GGTGAAGCTGGAGTAAAGTTTAC	833	

```

; PRIOR APPLICATION NUMBER: US/09/314,701
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(840)
US-10-314-639-23

```

Query Match	20.0%	Score 162.2	DB 3	Length 840
Best Local Similarity	53.8%	Pred. No. 9.4e-32		
Matches 432	Conservative 0	Mismatches 353	Indels 18	Gaps 4
Qy	22	GTAGGTTTACGCTGAGTACATTGTTTTTTTCTTATCTGATGGTCTTTTTCTGTATGCA	81	
Db	34	GTACTTGCACTCTTATATCATTTCTTATCTTATGGAATCCCTTTTCAGCTATAAATCATTAAT	93	
Qy	82	AAATTTTCTGAAGGAGGAGGAGGACTTTATATAGGTAGTCAGTATAAAGTTGGTATTCCTCC	141	
Db	94	CATACAGGAATTAACACTAGTGGTATATATATACAGGCGAGTATAGCAGGAGTATCC	153	
Qy	142	AAATTTTGTAAATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAGATTTTTTTCGG	201	
Db	154	CAATTTAGCAATTTCTCAGTAAAAGAACTAAATGTTGATACAATACAACTAGTAGGATAT	213	
Qy	202	TTAGGCTTTCATTAAGTCT--GAGATAAATACCTCACAGCAATTTTACAGCATCATATGAC	258	
Db	214	AAAAAAGTCGCTTCTATTCGATCCCTAACACTTATTCAAACTTTCAAGGTCCTATATCT	273	
Qy	259	CCTACTTATGCAAGCAGTTTTGCAAGGTTTTAGTGGTATCATTTGATATTTATGTTAAATGAC	318	
Db	274	GTTACATTTCAAGATAATGCTGCTAGTTTTCAGTGGAGCAATTTGATATTTCTTACCCGAA	333	
Qy	319	TTT---AGGGTAGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTAC	375	
Db	334	AGCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTGATGTCAAAGATCCTAAAGAC	393	
Qy	376	CCTGAGAATAGCCAAAGCTACAAATTTTTTGTCTTGTCTCGAATGCTACAAATAGTAT	435	
Db	394	TACTCAGCAAAAAGATGCTTTTAGGTTTTTTTGTCTAGCAGCAATAACGCTACTACTGTT	453	
Qy	436	AAT-----AAGTTTATAGTACTAGAGAAATAACGCGTTGTTGACAAGTCTCTTAAT	486	
Db	454	CCTGATGCTCAAAAATATACAGTTATGAAGAAATAATGGCTTATCTGTTGCAATCAATCATG	513	
Qy	487	GTAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTGCT	546	
Db	514	ATCAATGGTTGTATGATCTATCTTTTAAATTTAGTCGTATCACCTTATATATGTGCA	573	
Qy	547	GGTGTGGTGAGATATATAAAGTTTTTATAGGTATATACATTTGCCTAAGTTTTCTTATCAA	606	
Db	574	GGTATTGGTGAAGATTTCAITGAAATTTTTTGTATCTTTGACATCTTAACTTCTTATCAA	633	
Qy	607	GTTAAGTTTTGGTGCACTACCTCCTAAATGTTTAAATCTATGTTGTTGGTGGGGGTTAT	666	
Db	634	GGAAAACTAGGTATTTAGTTATTAAGTCTTCTTCCCTAAGATTAATGATTTTGGTGGGGTAC	693	
Qy	667	TACCATAAGGTTGTAGGTGATAGGCATGAGAGGTAGAAAAATAGCTTACCATCCTACTGCA	726	
Db	694	TATCATAGAGTTATAGGGAATAAATTTAAAAATTT--AAATGTTAACCATGTTGTTACA	750	
Qy	727	TTATCTGACGTTCTCTAGAACTACTTTCAGCTTCTCTACTTTTAAATACGTATTTTGGT	786	
Db	751	CTTGATGAAATTTCTCTAAAGCACTTCTGCGTAGCTACACCTTAATGTTGCTTATTTTGGT	810	
Qy	787	TGGGAGATTTGGATTTAGATTTGC	809	
Db	811	GGTGAAGCTGGAGTAAAGTTTAC	833	

```
RESULT 5
US-10-059-964A-23
; Sequence 23, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/059,964A
; CURRENT FILING DATE: 2002-01-28
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
US-10-059-964A-23

Query Match      20.0%; Score 162.2; DB 3; Length 840;
Best Local Similarity 53.8%; Pred. No. 9.4e-32;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY      22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      34 GTACTTGTCATCTATATATATCTTCTATTTGATCCCTTTCAGCTATAATCATAT 93

QY      82 AATTTTCTGAAGGAGGAGAGGACTTTATATAGTAGTACGATATAAAGTTGGTATCCC 141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      94 CATACAGGAATAAACACTAGTGTGTATATATATACAGGGCAGTATAGACAGGAGTATCC 153

QY      142 AATTTTGTAGTAATTTTTCAGCTGAGGAACAATTCCTGGTATTACAAAAGATTTTGGG 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      154 CATTTTGTAGCAATTTCTCAGTAAAGAAACTTAATGTTGTATACAACTAGTAGGATAT 213

QY      202 TTAGGTCTTGATAAGTCT---GAGATAAATACTCACAGCAATTTTACAGCATCATATGAC 258
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      214 AAAAAAGTGGCTCTTATCGATCCATACACTTATTCACATTTCAAGGTCATATACT 273

QY      259 CTTACTTATGCAAGCAGTTTTTGCAGGGTTTAGTGGTATCATTTGATATTTATGATGAC 318
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      274 GTTACATTTCAAGATAATGCTGCTAGTTTTCAGTGGAGCAATTTGGATATTTCTTACCCGAA 333

QY      319 TTT---AGGGTAGAATTTGAAAGTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTAC 375
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      334 AGCTAAGACTTGAACCTTGAAGTCTTACGAAAAATTTGATGTCAAAGATCCTTAAGAC 393

QY      376 CCTGAGAATAGCCAAAGCTACAAATTTTGTCTTCTCGAAATGCTACAAATAGTGAT 435
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      394 TACTCAGCAAAAGATGCTTTTAGTTTTTGTCTTAGCAGTAATAGTCTACTAGTT 453

QY      436 AAT-----AAGTTTATAGTACAGATAAATACGGGGTTGTGCAAGTCTCTTAAT 486
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      454 CCTGATGCTCAAAAATATACAGTTATGAAGATAATAGGCTTATCTGTTCATCAATCATG 513

QY      487 GTAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCT 546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      514 ATCAATGGTTGTATGATCTATCTTTTAATAATTTAGTCGTATCACCTTATATATGTCGA 573

QY      547 GGTGTTGGTGCAGATATATAAAGTTTTTGTAGGTATATCAATGCTTAAGTTTTCTTATCAA 606
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      574 GGTATTGGTGAAGATTTCAATGAAATTTTGTATCTTTCACATTAATTAATCTTATCAA 633

QY      607 GTTAAGTTTGGTGTCAACTACCCCTTAATAGTTTAATATATGTTTGGTGGGGGTAT 666
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      634 GGAATACTAGGTATTAGTTATTACTTTCTTCTTAAGATTAATGATATTTGCTGGGTGCTAC 693
```

```
QY      667 TACCATAAGGTTGTAGGTGATAGGCATGAGAGAGTAGAAATAGCTTACCATCTACTGCA 726
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      694 TATCATAGATTATAGGGAATAATTTAAATTT---AATGTTAAACATGTTGTTACA 750

QY      727 TTATCTGACGTTCCCTAGAACTACTTCAGCTTCTGCTACTTTTAAATPACTGATTTTGGT 786
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      751 CTTGATGAATTTCTTAAAGCAACTTCTGCAGTAGCTACACTTAATGTTGCTTATTTTGGT 810

QY      787 TGGGAGATTGGATTTAGATTGGC 809
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      811 GGTGAAGCTGGAGTAAAGTTTAC 833

RESULT 6
US-09-314-701-25
; Sequence 25, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, Yasuko
; APPLICANT: Ohashi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(852)
US-09-314-701-25
```

```
Query Match      18.5%; Score 150.8; DB 3; Length 852;
Best Local Similarity 54.6%; Pred. No. 7.5e-29;
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

QY      96 GAGGAGAGGACTTTTATATAGGTAGTACAGTATAAAGTTGGTATTCCTCAATTTTGTAAATTT 155
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      111 GTGCACCTGGCTATATGTGAGTGGACAATAAACCCTACTGTTTCTCACTTTAGTAAATTT 170

QY      156 TTCAGCTGAAGAAACAATTCCTGGTATTACAAAAGATTTTGGTGTAGGTCTTGATAA 215
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      171 TTCACCTTAAAGAAACCTTATCTGACACTAAAGAGTTTATAGGACTPAGCAAAAGATATTAA 230

QY      216 GTCT-----GAGATAAATACTCACAGCAATTTTACAGCATCATATGACCTTACTTTATGC 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      231 GTCTATTACAGATATACAAACAATAAATAATTCACATTCCTTATACACAAAATTTCA 290

QY      270 AAGCAGTTTTGCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTTT---AGGGT 326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      291 AGATAATGCTGTAGCTTTCAGTGCAGCTGTGGATATATTTCCCAAGACAGTCCAAGGCT 350

QY      327 AGAATTTGAAAGTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTACCCTGAGAAATAG 386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      351 TGAGTAGAATGGTCTTATGAAGAAATTTGACGTTAAAAATCTCGGTAATTTACGTAGTAGA 410

QY      387 CCAAAAGCTACAAAATTTTGTCTCGAAATGCTACAAATAGTGAT----- 435
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      411 TGAAGCTTCAGGTATATTTGCTTTAGCAAGAGGAATTTGATAATCTTCAAAAATATCCTGA 470

QY      436 ----AATAAGTTTATAGTACTAGAGAAATAACGGCTTGTGTGACAAAGTCTCTTTAATGTAAA 491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      471 AACAAATAAGTATGTTGTTATAAAGACAATAGGCTTATCTGTCGCATCCATTATAATCAA 530

QY      492 TGTGTTTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCTGGTGTCT 551
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      531 TGGCTGTTATGATTTTCTTTTAAACAATTTTAAAGATATCACCTTACATATGCGTAGGGT 590

QY      552 TGGTGCAGATTATATAAAGTTTTTTTAGGTATATCATTTGCTAAGTTTTTCTTATCAAGTTAA 611
```



Db 591 TGGTGGGCAATTAATAGATTTTATAGTCTGTAGTGTAAATTTTGCCTTATCAAGTAA 650  
Qy 612 GTTTGGTGTCAACTACCTCTAAATGTTAATACTATGTTGTTGGTGGGGTTATTACCA 671  
Db 651 GGTAGGTATCAGTTATCCATTTCTCTAATATGATTATATTTGCTGACGGATATTACCA 710  
Qy 672 TAAGTTGTAGGTGATAGGATGAGAGATAGAAATAGCTTACCATCTACTGCATATC 731  
Db 711 TAAGGTCATAGGAATAAATTTAAACAATTTAAATGTTCAACAGGTTGTTAGTCTTAACAG 770  
Qy 732 TGACGTTCTAGAACTACTTTCAGCTTCTGCTACTTTAAATAGTATTATTTGGTGGGA 791  
Db 771 TCA---TCCTAAGTCTACTTTTGCACTAGTACTCTTAATGTTGAGTATTCGGTAGTA 827  
Qy 792 GATTGGATTTAGATTT 807  
Db 828 ATTGGGTTAAATTT 843

## RESULT 7

US-10-314-639-25

; Sequence 25, Application US/10314639

; Patent No. 6893640

; GENERAL INFORMATION:

; APPLICANT: Rikihisa, Yasuko

; APPLICANT: Ohasi, No. 6893640io

; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia

; TITLE OF INVENTION: Chaffeensis

; FILE REFERENCE: 22727/04021

; CURRENT APPLICATION NUMBER: US/10/314,639

; CURRENT FILING DATE: 2002-12-09

; PRIOR APPLICATION NUMBER: US/09/314,701

; PRIOR FILING DATE: 1999-05-19

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 25

; LENGTH: 852

; TYPE: DNA

; ORGANISM: Ehrlichia chaffeensis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(852)

US-10-314-639-25

Query Match 18.5%; Score 150.8; DB 3; Length 852;  
Best Local Similarity 54.6%; Pred. No. 7.5e-29;  
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

Qy 96 GAGGAGGAGCTTTATATAGGTAGTCACTATAAAGTTGGTATCCCAATTTTAGTAATTT 155  
Db 111 GTGCACCTGGGCTATATGTCACTGAGCAATATAAACCTACTGTTCTCACTTTAGTAATTT 170  
Qy 156 TTCAGCTGAAGAACAAATTCCTGGTATTAACAAAAGATTTTGGTGTAGTCTTGATAA 215  
Db 171 TTCACCTTAAGAGAACTTATCTGACACTAAAGAGTTATTAGGACTAGCAAAAAGATATAA 230  
Qy 216 GTCT-----GAGATAAATACTCACAGCAATTTTACAGCATATATGACCCCTACTTATGC 269  
Db 231 GTCTATTACAGATATACACAATAAATAATTCACATTCCTTATTAACACAAATTTCA 290  
Qy 270 AAGCAGTTTTCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTT---AGGGT 326  
Db 291 AGATAATGCTGTAGCTTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGGT 350  
Qy 327 AGAATTGAGGTTCTTATAGAAATTTTGAACCTGAACACAAATGGTACCCCTGAGATAG 386  
Db 351 TGAGGTAGAAATGGTCTTATGAGAAATTTGACGTTAAAAAATCCTGGTAAATTAACGTAAG 410  
Qy 387 CCAAGCTACAAAATTTTGTCTCGAAATGCTACAAATAGTAT-----AGGGT 435  
Db 411 TGAAGCCCTCAGGTATATGCTTTTAGCAAGGAATTCATATCTTCAAAAATATCTCGA 470

Qy 436 ----AATAAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAACTCTCTTAATGTAAA 491  
Db 471 AACAAATAAGTATGTTGTTTATAAGAAACAATGGCTTATCTGTCGCAATCCATTATAATCAA 530  
Qy 492 TGTTTGTATGATATTGCTAGTGTAGTATTCCTTTAGCACCTTATATATGTCGCTGGTGT 551  
Db 531 TGCGCTGTATGATTTTCTTTAAACAATTTAAAGATATCACCTTTACATATGCGTAGGGTT 590  
Qy 552 TGGTGCAGATTATATAAAGTTTGTAGGTATATCATTTGCCCTAAGTTTCTTATCAAGTTAA 611  
Db 591 TGGTGGGACATTTATAGAAATTTTGTAGTCTGTGAAGTTTAAATTTGCTTTATCAAGTAA 650  
Qy 612 GTTTGGTGTCAACTACCTCTAAATGTTAATACTATGTTGTTGGTGGGGTTATTACCA 671  
Db 651 GGTAGGTATCAGTTATCCATTTCTCTAATATGATTATTTGCTGACGATATTACCA 710  
Qy 672 TAAGGTTGTAGGTGATAGGATGAGAGATAGAAATAGCTTACCATCTACTGCATATC 731  
Db 711 TAAGGTCATAGGAATAAATTTAAACAATTTAAATGTTCAACAGGTTGTTAGTCTTAACAG 770  
Qy 732 TGACGTTCTAGAACTACTTTCAGCTTCTGCTACTTTAAATAGTATTATTTGGTGGGA 791  
Db 771 TCA---TCCTAAGTCTACTTTTGCACTAGTACTCTTAATGTTGAGTATTTTCGGTAGTA 827  
Qy 792 GATTGGATTTAGATTT 807  
Db 828 ATTGGGTTAAATTT 843

## RESULT 8

US-10-059-964A-25

; Sequence 25, Application US/10059964A

; Patent No. 6923963

; GENERAL INFORMATION:

; APPLICANT: RIKIHISA, YASUKO

; APPLICANT: OHASHI, NORIO

; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia

; TITLE OF INVENTION: CHAFFEENSIS

; FILE REFERENCE: 22727-04109

; CURRENT APPLICATION NUMBER: US/10/059,964A

; CURRENT FILING DATE: 2002-01-28

; PRIOR APPLICATION NUMBER: 09/314,701

; PRIOR FILING DATE: 1999-05-19

; PRIOR APPLICATION NUMBER: 60/100,843

; PRIOR FILING DATE: 1998-09-18

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 25

; LENGTH: 852

; TYPE: DNA

; ORGANISM: Ehrlichia chaffeensis

US-10-059-964A-25

Query Match 18.5%; Score 150.8; DB 3; Length 852;  
Best Local Similarity 54.6%; Pred. No. 7.5e-29;  
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

Qy 96 GAGGAGGAGCTTTATATAGGTAGTCACTATAAAGTTGGTATTCCTCAATTTTAGTAATTT 155  
Db 111 GTGCACCTGGGCTATATGTCACTGAGCAATATAAACCTACTGTTCTCACTTTAGTAATTT 170  
Qy 156 TTCAGCTGAAGAACAAATTCCTGGTATTAACAAAAGATTTTGGCGTTAGTCTTGATAA 215  
Db 171 TTCACCTTAAGAGAACTTATACTGACACTAAAGAGTTATTAGGACTAGCAAAAAGATATAA 230  
Qy 216 GTCT-----GAGATAAATACTCACAGCAATTTTACAGCATATATGACCCCTACTTATGC 269  
Db 231 GTCTATTACAGATATACACAATAAATAATTCACATTCCTTATTAACACAAATTTCA 290  
Qy 270 AAGCAGTTTTCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTT---AGGGT 326  
Db 291 AGATAATGCTGTAGCTTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGGT 350

327	Qy	AGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGGAAGCAATGGTGACCTCGAGAAATAG	386
351	Db	TGAGGTAGAAATGCGTCTTATGAGAAGATTTGACGTTAAAAAATCCTGGTAATTTACGCTAGTAGAAG	410
387	Qy	CCAAAGCTACAAATTTTTTGTCTTCGAAATGCTACAAATAGTGAT-----	435
411	Db	TGAAGCGCTTCAGGTATATTGCTTTTAGCAAGAGGAATTTGATAATCTTCAAAAAATATCCTGA	470
436	Qy	----AATAAGTTTATAGTACTAGAGAAATAACGGCGTTGTTTGACAAGTCTCTTAATGTAAA	491
471	Db	AACAAATAAGTATGTTGTTTAAAGACAATGGCTTATCTGTGCGATCCAAATTAATATCAAA	530
492	Qy	TGTTGCTTATGATATTGCTAGTGGTAGTATTCTTCTTTAGCACCTTATATATGTGCTGGTGT	551
531	Db	TGGCTGTTATGATTTTCTTTTAAACAAATTTAAAGTATCACCTTACATATGCGTAGGGTT	590
552	Qy	TGCTGCAGATTAATATAAAGTTTTTATAGGTATATCAATTCGCTAAAGTTTTCTTATCAAGTTAA	611
591	Db	TGCTGGGGACATTTATAGAATTTTTTATGTCGTAAAGTTTTTAAATTTGCTTTATCAAGGTAA	650
612	Qy	GTTTGGTGTCAACTACCTCTAAAATGTTAAATACTATGTGTTGGTGGGGGTTTATTACCA	671
651	Db	GGTAGGTATCAGTTATCCAAATTTCTCTAAATGATTTATATTGCTGACGGATTTTATACCA	710
672	Qy	TAAAGTTGTAGGTGATAGGCATGAGAGAGTAGAATAAGCTTACCATCCTACTGCAATTATC	731
711	Db	TAAAGTCATAGGAATAAATTTAACAAATTTAAATGTTTCAACAGTTGTTAGTCTTAAACAG	770
732	Qy	TGACGTTCTCTAGAACCTACTTCACGTTCTCTGCTACTTTTAAATACTGATTTATTTGGTTGGGA	791
771	Db	TCA---TCCTAAGTCTACTTTTTCGAGTAGCTACTCTTAAATGTTGAGTATTTTCGGTAGTCGA	827
792	Qy	GATTGGAATTTAGATTT	807
828	Db	ATTTGGGTTAAAAATTT	843

```

RESULT 9
US-09-660-587-41
; Sequence 41, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/660,587
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 41
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-2
US-09-660-587-41

Query Match      18.1%; Score 147.2; DB 3; Length 840;
Best Local Similarity 53.2%; Pred. No. 6.2e-28;
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5

Qy      1  ATCAATTACAAAGATTGTGCTGAGCTGTACGCTGAGTACATTGTTTTTCTTCTATCT 60
          |||||
Db      1  ATCAATTATAGAAATCTAGTAAAGCGGTTATCTCATTAATGTCATCTTACCA 60
          |||||

Qy     61  GATGGTGCTTTTCTGA-----TGCAAATTTTCTGAAGGAGGAGGACTTTAT 111
          |||||

Db     61  TATCAGCTCTTTGCGACATCCTGTAGGTTCAAGAACTAATGATTAACAAGAGCGCTTCTAC 120
          |||||

```

Qy	112	ATAGGTAGTCAGTATAAAGTTGGTATTCCTCCAAATTTTCTAGTAATTTTTCAGCTGAAGAAACA	171
Db	121	ATTAGTGCAAGATCAATCCAAAGTATATACACATTTTGAATAATCTCTGCTGAAGAAACT	180
Qy	172	ATTCTCTGGTATTACAAAAGAGATTTTTCGGTGTAGGCTCTTG-----ATAAGTCTGAG	222
Db	181	CCTATTAAATGGAAACAAATTTCTCTCACATAAAAAAGTTTTCGGAGTAAAGAAAGATGGTGAT	240
Qy	223	ATAAATACTCACAGCAATTTTACACGATCATATGACCCCTACTTAT-----GCAAGCAGT	276
Db	241	ATACAAAAAGAGCAGATTTTACAGAGTAGCTCCAGGCAATGATTTTCAAAATAACTTA	300
Qy	277	TTTGCAGGGTTTACTGGTATCATTTGGATATTATGTTAATAGCTTTTGGGTAGAAATTTGAA	336
Db	301	ATATCAGGATTTTCAGGAAGATTTGGTTACTCTATGGACGGACCAAGAAATAGAACTTTGAA	360
Qy	337	GGTTCATTATGAGATTTTGAACCTGAAAGACAAATGGGTACC-----CTGAGNAATAGCCAA	390
Db	361	GCTGCAATATCAACAATTTAAATCCAAAAAACAACCGATAACAATGATPACTGATTAATCGTGAA	420
Qy	391	AGCTACAAAATTTTTCCTTCGTAATGCTACAAATAGTGATAAATAAGTTTATATGTA	450
Db	421	TACTATAAACATTTTTCGATTCCTGTAAGATGCAATGGGAAGATCAGCAATATGTAGTA	480
Qy	451	CTAGAGAATAACCGCGTTGTGTGACAAGTCTCTTAATGTAAATTTTGTGTATGATATTGCT	510
Db	481	CTTAAAAATGACGGCATAACTTTTATGTGCTAATGATGGTTAATACTTGTCTATGACATTACA	540
Qy	511	AGTGGTAGTATTCCTTTAGCACCTTATATGTGCTGCTGTGTGGTGCAGATTATATAAG	570
Db	541	GCTGAAGGAGTATCTTTCGTAACCATATGTCATGTGCAGGTATAGGAGCAGATCTTATCACT	600
Qy	571	TTTTTTAGGTATATCATTTGCCCTAAGTTTCTTATCAAAGTTAAAGTTTGGTGTCAACTACCCCT	630
Db	601	ATTTTTAAGACCTCAATCTAATAATTTGCTTACCAAGGAAATAAGGTATTAGTTACCCCT	660
Qy	631	CTAAATGTTAATACTATGTTGTTGGTGGGGTTATTAACATAAGGTTGTAGGTGATAGG	690
Db	661	ATCACACCAGAAAGTCTCTGCAATTTTATGGTGGATACTACCATGGCGTTTATTGGTAATAAA	720
Qy	691	CATGAGAGGTAGAAATAGCTTACCATCTACTGCATTTATCTGACGTTTCTAGAACTACT	750
Db	721	TTTGAGAAGATACCTGTA---ATAACTCTGTAGTATTAAATGATGCTCCTCAAAACCCACA	777
Qy	751	TCAGCTTCTGCTACTTTAAATPACTGATTATTTTGGTTCGGAGATTTGGAATTTAGATT	806
Db	778	TCGCTTCAGTAACTCTTGACGTTGGATCTTTTGGCGGAGAAATTTGGATGAGGTT	833

RESULT 10

US-09-811-007A-41

; Sequence 41, Application US/09811007A

; Patent No. 6660269

; GENERAL INFORMATION:

; APPLICANT: Walker, David H.

; APPLICANT: McBride, Jere W.

RESULT 10  
US-09-811-007A-41  
; Sequence 41, Application US/09811007A  
; Patent No. 6660269  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/811,007A  
; CURRENT FILING DATE: 2001/10-23  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46

```

1  APPLICANT:  McBridge, Jere W.
2  APPLICANT:  Yu, Xue-Jie
3  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
4  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
5  FILE REFERENCE:  D6152CIP2
6  CURRENT APPLICATION NUMBER:  US/09/811,007A
7  CURRENT FILING DATE:  2001-10-23
8  PRIOR APPLICATION NUMBER:  09/660,587
9  PRIOR FILING DATE:  2000-09-12
10  NUMBER OF SEQ ID NOS:  46
11  SEQ ID NO 41
12  LENGTH:  840
13  TYPE:  DNA
14  ORGANISM:  Ehrlichia canis
15  FEATURE:
16  OTHER INFORMATION:  nucleic acid sequence of E. canis p28-2
17  US-09-811-007A-41

```

```
Query Match      18.1%; Score 147.2; DB 3; Length 840;
Best Local Similarity 53.2%; Pred. No. 6.2e-28;
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;

QY 1 ATGAATTTACAAAAGATTGTTGTAGGTGTTACGCTGAGTACATTGTTTTTCTTATCT 60
DB 1 ATGNAATTAAAGAAATCTAGTAAGAAGCGGTTAATCTCATTAATGTCATCTTACCA 60

QY 61 GATGGTCTTTTCTGA-----TGCAAAATTTTCTGAAGGGAGGAGGACTTTAT 111
DB 61 TATCAGTCTTTTGCAGATCTGTAGGTTCAAGAACTAATGATAACAAGAAGGCTTCTAC 120

QY 112 ATAGGTAGTCAGTATAAAGTTGGTATTCCCAATTTTAGTAATTTTTCAGCTGAAGAAACA 171
DB 121 ATTAGTGCRAAGTACAATCCAAGTATATCACCTTTAGAAAATTTCTCTGCTGAAGAACT 180

QY 172 ATTCCTGGTATTACAAAAGATTTTTCGCTAGGCTCTG-----ATAAGTCTGAG 222
DB 181 CCTATTAAATGGAACAAATCTCTCACTAAAGAAAGTTTCGACTAAAGAAAGATGGTAT 240

QY 223 ATAAATCTCACAGCAATTTTACAGCATATATGACCTTACTTAT-----GCAAGCAGT 276
DB 241 ATACAAAAAAGACGATTTTACAAGTAGTCTCCAGGCATTGATTTTCAAAATAACTTA 300

QY 277 TTTGCGGGTTTAGTGATCATTTGATGATATATGATGATGATGATGATGATGATGATGAT 336
DB 301 ATATCAGGATTTTTCAGGAAGTATTTGATGATGATGATGATGATGATGATGATGATGAT 360

QY 337 GGTTCCTTATGAGAAATTTTGAACCTGAAAGCAATGGTACC-----CTGAGAATAGCCAA 390
DB 361 GCTGCATATCAACAATTTTAAATCCAAAAAACACCGATAAATGATGATGATGATGATGAT 420

QY 391 AGCTACAAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
DB 421 TACTATAAACAATTTTGCATTTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480

QY 451 CTAGAGAATAACGCGTGTGTTGACAAGTCTCTTAAATGTAATGTTTCTTATGATGATGATGAT 510
DB 481 CTTAAAAAATGACGCGCATTAATTTTATGTCATGATGATGATGATGATGATGATGATGATGAT 540

QY 511 AGTGGTAGTATTCCTTTTAGCACCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 541 GCTGAAGGAGTATCTTTTCGTAACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 571 TTTTATAGGTATATCATTTGCTTAAAGTTTCTTATCAAGTTAAGTTTGGTGTCAACTACCT 630
DB 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTACCAAGGAAAAATAGGTATTTAGTTACCT 660

QY 631 CTAATGTTAATACTATGTTGTTGGGGGTTTATACCAATAGGTTGATGATGATGATGATGATGAT 690
DB 661 ATCACCAGGAAGTCTCTGCAATTTATTTGGGGGATCTTACCATGCGGTTTATGGTAATAA 720

QY 691 CATGAGAGAGTAGAATAATAGCTTACCATCTCTACTGCAATTTCTGACGTTTCTCTAGAACTACT 750
DB 721 TTTGAGNAGATACCTGTA---ATAACTCTCTGTAGTATTTAATGATGCTCTCTCAACCCACA 777

QY 751 TCAGCTTCTGCTACTTTAAATACTGATTTATTTGGTGGGAGATTGGATTTAGATT 806
DB 778 TCTGCTTCAGTAACCTCTTGACGTTGGATACTTTGGCGGAGAAATTTGGNATGAGGTT 833
```

## RESULT 11

```
US-09-314-701-47
; Sequence 47, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: Rikhisia, Yasuko
; APPLICANT: Ohasi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chafteensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
```

```
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(843)
US-09-314-701-47
```

```
Query Match      18.1%; Score 147.2; DB 3; Length 843;
Best Local Similarity 53.2%; Pred. No. 6.2e-28;
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;
```

```
QY 1 ATGAATTTACAAAAGATTGTTGTAGGTGTTAGCTGAGTACATTGTTTTTCTTATCT 60
DB 1 ATGNAATTAAAGAAATTTCTAGTAAGAAGCGGTTAATCTCATTAATGTCATCTTACCA 60

QY 61 GATGGTCTTTTCTGA-----TGCAAAATTTTCTGAAGGGAGGAGGACTTTAT 111
DB 61 TATCAGTCTTTTGCAGATCTGTAGGTTCAAGAACTAATGATAACAAGAAGGCTTCTAC 120

QY 112 ATAGGTAGTCAGTATAAAGTTGGTATTCCCAATTTTAGTAATTTTTCAGCTGAAGAAACA 171
DB 121 ATTAGTGCRAAGTACAATCCAAGTATATCACCTTTAGAAAATTTCTCTGCTGAAGAACT 180

QY 172 ATTCCTGGTATTACAAAAGATTTTTCGCTAGGCTCTG-----ATAAGTCTGAG 222
DB 181 CCTATTAAATGGAACAAATCTCTCACTAAAGAAAGTTTCGACTAAAGAAAGATGGTAT 240

QY 223 ATAAATCTCACAGCAATTTTACAGCATATATGACCTTACTTAT-----GCAAGCAGT 276
DB 241 ATACAAAAAAGACGATTTTACAAGTAGTCTCCAGGCATTGATTTTCAAAATAACTTA 300

QY 277 TTTGCGGGTTTAGTGATCATTTGATGATATATGATGATGATGATGATGATGATGATGAT 336
DB 301 ATATCAGGATTTTTCAGGAAGTATTTGATGATGATGATGATGATGATGATGATGATGAT 360

QY 337 GGTTCCTTATGAGAAATTTTGAACCTGAAAGCAATGGTACC-----CTGAGAATAGCCAA 390
DB 361 GCTGCATATCAACAATTTTAAATCCAAAAAACACCGATAAATGATGATGATGATGATGAT 420

QY 391 AGCTACAAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
DB 421 TACTATAAACAATTTTGCATTTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480

QY 451 CTAGAGAATAACGCGTGTGTTGACAAGTCTCTTAAATGTAATGTTTCTTATGATGATGATGAT 510
DB 481 CTTAAAAAATGACGCGCATTAATTTTATGTCATGATGATGATGATGATGATGATGATGATGAT 540

QY 511 AGTGGTAGTATTCCTTTTAGCACCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 541 GCTGAAGGAGTATCTTTTCGTAACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 571 TTTTATAGGTATATCATTTGCTTAAAGTTTCTTATCAAGTTAAGTTTGGTGTCAACTACCT 630
DB 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTACCAAGGAAAAATAGGTATTTAGTTACCT 660

QY 631 CTAATGTTAATACTATGTTGTTGGGGGTTTATACCAATAGGTTGATGATGATGATGATGATGAT 690
DB 661 ATCACCAGGAAGTCTCTGCAATTTATTTGGGGGATCTTACCATGCGGTTTATGGTAATAA 720

QY 691 CATGAGAGAGTAGAATAATAGCTTACCATCTCTACTGCAATTTCTGACGTTTCTCTAGAACTACT 750
DB 721 TTTGAGNAGATACCTGTA---ATAACTCTCTGTAGTATTTAATGATGCTCTCTCAACCCACA 777

QY 751 TCAGCTTCTGCTACTTTAAATACTGATTTATTTGGTGGGAGATTGGATTTAGATT 806
DB 778 TCTGCTTCAGTAACCTCTTGACGTTGGATACTTTGGCGGAGAAATTTGGNATGAGGTT 833
```

## RESULT 12

[illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 28, 2005, 10:46:11 ; Search time 4515 Seconds  
(without alignments)  
10235.598 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaagaattgt.....ttggatttagattggccta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_mtg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	813	100.0	813	6	AR437336	AR437336 Sequence
C 2	813	100.0	11329	1	AF082744	AF082744 Ehrlichia
C 3	811.4	99.8	110000	1	CP000107_12	Continuation (13 o
C 4	809.8	99.6	28254	1	AF078553	AF078553 Ehrlichia
C 5	574.6	70.7	2977	1	AF068234	AF068234 Ehrlichia
C 6	574.6	70.7	27190	1	ECU72291	U72291 Ehrlichia c
C 7	386.6	47.6	21115	1	AY652746	AY652746 Ehrlichia
C 8	386.6	47.6	99920	1	CR925677_14	Continuation (15 o
C 9	380.6	46.8	24993	1	AV343331	AV343331 Ehrlichia
C 10	380.6	46.8	110000	1	CR767821_14	Continuation (15 o
C 11	380.6	46.8	110000	1	CR925678_14	Continuation (15 o
C 12	365.6	45.0	19449	1	AF319940	AF319940 Ehrlichia
C 13	162.2	20.0	840	6	AR303095	AR303095 Sequence
C 14	162.2	20.0	840	6	AR655791	AR655791 Sequence
C 15	162.2	20.0	27190	1	ECU72291	U72291 Ehrlichia c
C 16	160.6	19.8	1280	1	AF393389	AF393389 Ehrlichia
C 17	160.6	19.8	14759	1	AF230842	AF230842 Ehrlichia
C 18	155.6	19.1	19449	1	AF319940	AF319940 Ehrlichia

19 150.8 18.5 852 6 AR303096  
20 150.8 18.5 852 6 AR655792 Sequence  
21 150.8 18.5 21115 1 AY652746 Ehrlichia  
22 150.8 18.5 99920 1 CR925677\_14  
23 150.2 18.5 674 1 AB178801  
24 148.8 18.3 674 1 AB178799  
25 148.8 18.3 674 1 AB178800  
26 148.6 18.3 674 1 AB178794  
27 148 18.2 28254 1 AF078553 Ehrlichia  
28 148 18.2 110000 1 CP000107\_12  
29 147.8 18.2 674 1 AB178803  
30 147.6 18.2 5828 1 AF479838  
31 147.2 18.1 674 1 AB178802  
32 147.2 18.1 840 6 AR437334  
33 147.2 18.1 843 6 AR303107  
34 147.2 18.1 843 6 AR655803  
35 147.2 18.1 906 1 AF528511  
36 147.2 18.1 906 1 AF528512  
37 147.2 18.1 906 1 AF528514  
38 147.2 18.1 906 1 AF528515  
39 147.2 18.1 6913 1 AF324792  
40 147.2 18.1 11329 1 AF082744  
41 147.2 18.1 110000 1 CP000107\_11  
42 147 18.1 674 1 AB178798  
43 146.6 18.0 726 6 AR303114  
44 146.6 18.0 726 6 AR655810  
45 146.6 18.0 3507 1 AF125276

ALIGNMENTS

RESULT 1  
AR437336  
LOCUS AR437336 813 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 45 from patent US 6660269.  
ACCESSION AR437336  
VERSION AR437336.1 GI:40202248  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 813)  
AUTHORS Walker,D.H., Yu,X.-J. and McBride,J.W.  
TITLE Homologous 28-kilodalton immunodominant protein genes of Ehrlichia canis and uses thereof  
JOURNAL Patent: US 6660269-A 45 09-DEC-2003;  
Research Development Foundation; Carson City, NV  
FEATURES  
source  
1..813  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 813; DB 6; Length 813;  
Best Local Similarity 100.0%; Pred. No. 1.3e-137; Indels 0; Gaps 0;  
Matches 813; Conservative 0; Mismatches 0;  
Qy 1 ATGAATTACAAAAGATTGTTGTTAGGTGTTACGCTGAGTACATTGTTTTCTTATCT 60  
Db 1 ATGAATTACAAAAGATTGTTGTTAGGTGTTACGCTGAGTACATTGTTTTCTTATCT 60  
Qy 61 GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACCTTTATATAGGTAGT 120  
Db 61 GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACCTTTATATAGGTAGT 120  
Qy 121 CAGTATAAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT 180  
Db 121 CAGTATAAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT 180  
Qy 181 ATTACAAAAGAGTTTTTGGGTTAGGCTTTGATAAGTCTGAGATAAATACTCACAGCAAT 240  
Db 181 ATTACAAAAGAGTTTTTGGGTTAGGCTTTGATAAGTCTGAGATAAATACTCACAGCAAT 240









QY	721	ACTGCATTCTCAGCGTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTTAAATACTGATTAT	780
Db	10126	ACTGCATTCTCAGCGTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTTAAATACTGATTAT	10067
QY	781	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	813
Db	10066	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	10034
RESULT 3			
CP000107_12/c			
WPCOMMENT			
Sequence split into 14 fragments LOCUS CP000107 Accession CP000107			
Fragment Name Begin End			
CP000107_00 1 110000			
CP000107_01 100001 210000			
CP000107_02 200001 310000			
CP000107_03 300001 410000			
CP000107_04 400001 510000			
CP000107_05 500001 610000			
CP000107_06 600001 710000			
CP000107_07 700001 810000			
CP000107_08 800001 910000			
CP000107_09 900001 1010000			
CP000107_10 1000001 1110000			
CP000107_11 1100001 1210000			
CP000107_12 1200001 1310000			
CP000107_13 1300001 1315030			
Continuation (13 of 14) of CP000107 from base 1200001 (CP000107 Ehrlichia canis str. Jak			
Query Match 99.8%; Score 811.4; DB 1; Length 110000;			
Best Local Similarity 99.9%; Pred. No. 4.6e-138;			
Matches 812; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGAATTACAAAAGATTGTGTAGGTGTGTACGCTGAGTACATTTGTTTTTTCTTATCT	60
Db	83345	ATGAATTACAAAAGATTGTGTAGGTGTGTACGCTGAGTACATTTGTTTTTTCTTATCT	83286
QY	61	GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGACTTTATATAGTAGT	120
Db	83285	GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGACTTTATATAGTAGT	83226
QY	121	CAGTATAAAGTGTGTAATCCCAATTTAGTAAATTTTTCAGCTCAAGAAACAATTCCTGCT	180
Db	83225	CAGTATAAAGTGTGTAATCCCAATTTAGTAAATTTTTCAGCTCAAGAAACAATTCCTGCT	83166
QY	181	ATTACAAAAGATTTTTTGGCTTAGGTCTTTGATAAGTCTGAGATAAATACTCACAGCAAT	240
Db	83165	ATTACAAAAGATTTTTTGGCTTAGGTCTTTGATAAGTCTGAGATAAATACTCACAGCAAT	83106
QY	241	TTTACAGCATCATAGCCCTACTTATGCAAGCAGTTTTCAGCGGTTTAGTGTATCATTT	300
Db	83105	TTTACAGCATCATAGCCCTACTTATGCAAGCAGTTTTCAGCGGTTTAGTGTATCATTT	83046
QY	301	GGATATTATCTTAATGACTTTTAGGTAGAAATTTGAAGTCTTATGAGAATTTTGAACCT	360
Db	83045	GGATATTATCTTAATGACTTTTAGGTAGAAATTTGAAGTCTTATGAGAATTTTGAACCT	82986
QY	361	GAAGAACAATGGTACCCCTGAGAATAGCAAGCTACAAAATTTTTTGTCTCGAAAT	420
Db	82985	GAAGAACAATGGTACCCCTGAGAATAGCAAGCTACAAAATTTTTTGTCTCGAAAT	82926
QY	421	GCTACAAATAGTATTAATAGTTTATAGTACTAGAGATAACCGGTTTGTGAAGTCT	480
Db	82925	GCTACAAATAGTATTAATAGTTTATAGTACTAGAGATAACCGGTTTGTGAAGTCT	82866
QY	481	CTTAATCTAAATGTTTGTATGATATTTAGTGTAGTATTTCTTTAGCACCTTATATG	540
Db	82865	CTTAATCTAAATGTTTGTATGATATTTAGTGTAGTATTTCTTTAGCACCTTATATG	82806
QY	541	TGTGCTGGTGTGGTGAGATTATATAAAGTTTTTAGGTATATCATTTGCCAAAGTTTTCT	600
Db			

Db	82805	TGTGCTGGTGTGGTGAGATTATATAAAGTTTTTAGGTATATCATTTGCCAAAGTTTTCT	82746
QY	601	TATCAAGTTAAGTTTGGTGTCAACTACCTCTAAATGTAACTATCTTTTGGTGG	660
Db	82745	TATCAAGTTAAGTTTGGTGTCAACTACCTCTAAATGTAACTATCTTTTGGTGG	82686
QY	661	GGTATTATACCATTAAGTTGTAGGTGATAGGATGAGAGTAGAATAAGCTTACCATCT	720
Db	82685	GGTATTATACCATTAAGTTGTAGGTGATAGGATGAGAGTAGAATAAGCTTACCATCT	82626
QY	721	ACTGCATTATCTGACGTTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTAAATACTGATTAT	780
Db	82625	ACTGCATTATCTGACGTTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTAAATACTGATTAT	82566
QY	781	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	813
Db	82565	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	82533
RESULT 4			
AF078553/c			
LOCUS			
DEFINITION			
1, complete sequence.			
ACCESSION			
AF078553 AF078554 AF078555 AH006958			
VERSION			
AF078553.2 GI:13512584			
KEYWORDS			
SOURCE			
ORGANISM			
Ehrlichia canis			
Ehrlichia canis			
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;			
Anaplasmataceae; Ehrlichia.			
REFERENCE			
1 (bases 1 to 28254)			
Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.			
Cloning and characterization of multigenes encoding the			
immunodominant 30-kilodalton major outer membrane proteins of			
Ehrlichia canis and application of the recombinant protein for			
serodiagnosis			
J. Clin. Microbiol. 36 (9), 2671-2680 (1998)			
JOURNAL			
PUBMED			
9705412			
REFERENCE			
2 (bases 1 to 28254)			
Ohashi,N., Rikihisa,Y. and Unver,A.			
Analysis of transcriptionally active gene clusters of major outer			
membrane protein multigene family in Ehrlichia canis and E.			
chaffeensis			
JOURNAL			
PUBMED			
11254561			
REFERENCE			
3 (bases 1 to 28254)			
Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.			
Direct Submission			
Submitted (16-JUL-1998) Department of Veterinary Biosciences, The			
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA			
JOURNAL			
REFERENCE			
4 (bases 1 to 28254)			
Ohashi,N., Rikihisa,Y. and Unver,A.			
Direct Submission			
Submitted (29-NOV-2000) Department of Veterinary Biosciences, The			
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA			
JOURNAL			
AUTHORS			
TITLE			
REMARK			
COMMENT			
On or before Apr 2, 2001 this sequence version replaced gi:3790556,			
gi:3790555, gi:3790558, gi:3790557.			
FEATURES			
source			
1..28254			
/organism="Ehrlichia canis"			
/mol_type="genomic DNA"			
/strain="Okiahoma"			
/db_xref="taxon:944"			
<1..615			
/codon_start=1			
/transl_table=11			
/product="hypothetical transcriptional regulator"			
/protein_id="AAK28679.1"			
/db_xref="GI:13512585"			
/translation="PEYANEIKAHPLIEDLEIKNIQHKFTGEGIRLDVDDYASKNF			
KKEGAESTKELKSAATKVRPHFPVDECVGKEIKRQIRMRGMSQNLANKLGITTFQOVQKY			

```

EGKTRIVISRLVQLASLVNVEVRDMLKLQEDLKNISCDNPITPPHALRDNEKEFLP
EFNDSKIDSKEVLMMVRAYTCIRKEKVRNIYINLVKALSNDKS"
843. .1733
/gene="p30-19"
CDS
843. .1733
/gene="p30-19"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-19"
/protein_id="AAK28680.1"
/db_xref="GI:13512586"
/translation="MKLLYHLDMIMIKESAIGIVSFIALFAPNAPSPVPIDFSNES
EMAGFYASAQINIGPFPSPISAKIYTDKSEKELTLFSKEETETIDLKAGDFKRG
YSPYNNRYTGFGAIGYGGGLRGVELEGGTFPDVDKQKPNPDGHRYPALSKDSEI
QNSGSSSSKNDKVVNMKEGFMNAISFMFNACDITNIGSSLVFNACIGIQGAIIRFL
GCTNIHTLFAKGLGLGLISPKTILFANGYVYKADNAFTNLSVQYVPEISAAKPHID
PIVYFNADNYGCEVGLRFL"
2491. .3318
/gene="p30-18"
CDS
2491. .3318
/gene="p30-18"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-18"
/protein_id="AAK28681.1"
/db_xref="GI:13512587"
/translation="MNNKLSLLYALILFTSHVSSALVLDHNLVYFGIOYKPARHLL
SNLLIKESQDVVEVLALKYDAIGSPLDSTKEVNNFTIKYDPHVDNNRGLFSGVY
YNKFRIRSEISHFQIKNEGHRKRGFEKFAKLPAPSTQGYRHVTLNNGISTY
SALINACYDLVLPANHITITCYCLGFGIDVIDFLSKYTKTFKSHQKLGASYPISHRMSV
FTEVYHGLFKGFEQLPLNANTSPPOPPHVHTTASAILSIGYGGSVGIKFIL"
3425. .3871
/codon_start=1
/transl_table=11
/product="unknown function U2"
/protein_id="AAK28682.1"
/db_xref="GI:13512588"
/translation="MNNKLVQNEVLLLLALPYSLHNSIDIEYNEQPIKLSRLMKNIT
KSQIKHNYKYLKLLKKISIGWLFKFSNDYSNOVIYLRSLSKLVSTIALFDPK
NIFLVYKVLDSLNRFSNANYFYSKHNTITNCFVSRAYFPLSCSY"
3882. .4832
/gene="p30-17"
CDS
3882. .4832
/gene="p30-17"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-17"
/protein_id="AAK28683.1"
/db_xref="GI:13512589"
/translation="MNYAKVFLVCILIFPLSPFSFATNNYFVHEIGKSGHFGYVQ
YKGPTRPNRSIADDSFTNLLALSHKDYKDLFVSSTVRGLFSLPQSQOHLHVAATGG
STLTLNLDKSNKFIIPGNYPTIDNLGVGGIVGYSINNRLEAEFAFEKKNIKAPTYG
NYDEYFAIATVYVKGKTPVHYCKMNTGILSSFLVNTCYDFTLAKKIAIAPYCL
GVGDFIDFLQOTRLKASQYAKAGLSAISPNLTFVVDGSPHGMNMQPGLLVDPYT
DISVMSPGDNATAYSEBTTMLAKLNMIFLAGSIGIRFIS"
4848. .5699
/gene="p30-16"
CDS
4848. .5699
/gene="p30-16"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-16"
/protein_id="AAK28684.1"
/db_xref="GI:13512590"
/translation="MYKLYLSFIISLAQLLFGSFAFSDIKNNNIHSGYITIKYQPTI
SNFKNFKITDFTDPTDPIGDIITAPNTNPFDLKNYNFSLVYHKDSYKFPENDYSLG
ALSGTLAKVNIIRIEPESYKNPDTKRLAYHRSRGHKFFAIPRTSNFGVLPENDYTV
ANNNGISIIISNIIINLCSTKTKNFTPYICLIGIGDFIETFDVMRKVTKYQKGVISYP
ITPKLVLSISQYHKVIGNKPKFLPLIQFVALKRTDSDNSPEDKDVLTALLTLDLEHFSSE
IGLSFIF"
5716. .6510
/gene="p30-15"
CDS
5716. .6510
/gene="p30-15"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-15"
/protein_id="AAK28685.1"
/db_xref="GI:13512591"
/translation="MNYNIYVIKYTIALAFLLPRVSPSIIIGNIEKSIKLLSVHIN
SOYKPSISOISNVLQIENNSEKKINILNSNTTYNNQLENSITNPFPIIGYFKR
LRPAVEDSEYEBHIKDNDSLKANLSKYKMYNEDQNFETIATDNKLSITSATVNICY
LDLNNITVPHLCTAVGICSTGFFNDMPFKLLYQKRGIGLYLINNSVMLFPVYVHK
VMRNKLKLLTQYSDINAFDLDAITVLANTDIDIGYFGEVGVRFIN"
6525. .7394
/gene="p30-14"
CDS
6525. .7394
/gene="p30-14"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-14"
/protein_id="AAK28686.1"
/db_xref="GI:13512592"
/translation="MLQRLNFINIILAFLLLPFPOSFTLYIHDHEITQNVGLYISQO
YKPSIYFNFELIEENSHKTVELMGLANDVTHVTEYVLKDNKTFTPYSAKFRNSLIN
LSGAIYSGQQPRLIEIGSEYENFDVASCNCPVKANRYIALVRDKKPGNIYPODHS
HNSMYVTIKNNGISILSWMINGCYDIAFSNVKISPYCAGIGGDFITLFTWHIKF
AVQKGEIYLSVSPSISIFANGHYHKVMDNVFKNLHVKIYIQLQDAPITTSARAKLRI
GTFGEVGVRFV"
7419. .8255
/gene="p30-13"
CDS
7419. .8255
/gene="p30-13"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-13"
/protein_id="AAK28687.1"
/db_xref="GI:13512593"
/translation="MNNKSLILGTLISLFSLLPIKAFSVINSHDISNNVNGLYFTG
QRPVSHFSGFTVRETNIATQQLVSLNTKNKNHIIERTNPSGIYTAKEQDNAASES
GAIGYSYPEGLKPEIEISYKFGVGRKNTQSNVAFALARTQTSNPSDNKYVVMK
NSGLSVASYMINGCNMSPNVLVSPYICAGIGEDFIEFTDLYIKLAYQKGLGVNYS
LSSRFNIFADMYHKYVIGNQKLNLIHVALDTPPKVTSAIATLNVAYPGGEVGIRF
IL"
8266. .8823
/codon_start=1
/transl_table=11
/product="unknown function U3"
/protein_id="AAK28688.1"
/db_xref="GI:13512594"
/translation="MATYMHAKAIIEPMESPILTDYIILSLNNVODHNSFTICIT
KLYKLISSTIKKILINVKONIAVHLNHLINLTLYNKKSKFSTQSFYEMFYTKINOR
ILNFNLSYFTKLVNTQKNESYKSYTKYSHNFNATYTTTSFIHQHKLTKYSITNOR
NKSTKSPSKILAKNLLNNQENSEY"
8816. .9685
/gene="p30-12"
CDS
8816. .9685
/gene="p30-12"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-12"
/protein_id="AAK28689.1"
/db_xref="GI:13512595"
/translation="MNSKTFSLIGLSILICLAACLPQISFSESBNVTYNTWHTGLYIS
GLKPSYFSDFSFIKEFTYNTNTEALFGLKQDISILRNKETTQYNNNFYPTAKFOD
DRASFJAVGYIANNGPRIIEIGSEYEFVKNPGNYTIDAHRYIALAREKTSYLLS
PKNKVIYKNGGISIVSIIINGCYDISLSDSKVSPYICTGFGDFTETFSAIRFKA
YQKIGISYSLSSNIIIFTDGYVHKVINSQPKNLNVEHVNELTTDPKVTSATAFLNI
```

gene		EYPGFGFLKFIIP" 9694. .10533		ORGANISM		Ehrlichia chaffeensis	
Query Match		99.6%; Score 809.8; DB 1; Length 28254;		REFERENCE		Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;	
Best Local Similarity		99.8%; Pred. No. 1.4e-137;		AUTHORS		Anaplasmataceae; Ehrlichia.	
Matches 811; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				TITLE		1 (bases 1 to 2977)	
QY		1 ATGAATTTACAAAAGATTGTTGTGAGGTGTACGCTCAGTACATTTCTGTTTTTCTTATCT 60		JOURNAL		Yu,X.J., McBride,J.W. and Walker,D.H.	
DB		24807 ATGAATTTACAAAAGATTGTTGTGAGGTGTACGCTCAGTACATTTCTTATCT 24748		REFERENCE		Genetic diversity of the 28-kilodalton outer membrane protein gene in human isolates of Ehrlichia chaffeensis	
QY		61 GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGACTTTATATAGTACT 120		JOURNAL		J. Clin. Microbiol. 37 (4), 1137-1143 (1999)	
DB		24747 GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGACTTTATATAGTACT 24688		REFERENCE		10074538	
QY		121 CAGTATAAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGTT 180		AUTHORS		2 (bases 1 to 2977)	
DB		24687 CAGTATAAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGTT 24628		JOURNAL		Yu,X.J., McBride,J.W. and Walker,D.H.	
QY		181 ATTACAAAAAGATTGTTGGTTAGGTCTTGATAAGTCTGAGATAAATACTCACGCAAT 240		REFERENCE		Direct Submission	
DB		24627 ATTACAAAAAGATTGTTGGTTAGGTCTTGATAAGTCTGAGATAAATACTCACGCAAT 24568		AUTHORS		Submitted (26-MAY-1998) Pathology, The University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA	
QY		241 TTTACACGATCATATGACCTACTTATGCAAGCAGTTTTTCAGGGTTTGTGTTATCATTT 300		JOURNAL		4 (bases 1 to 2977)	
DB		24567 TTTACACGATCATATGACCTACTTATGCAAGCAGTTTTTCAGGGTTTGTGTTATCATTT 24508		REMARK		Yu,X.J., McBride,J.W., Zhang,X. and Walker,D.H.	
QY		301 GGATATATGTTAATGACTTTAGGTGAGAAATTTGAAGGTTCTTATGAGAAATTTGAACCT 360		FEATURES		Direct Submission	
DB		24507 GGATATATGTTAATGACTTTAGGTGAGAAATTTGAAGGTTCTTATGAGAAATTTGAACCT 24448		source		Submitted (01-FEB-2000) Pathology, University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA	
QY		361 GAAAGACAATGGTACCTGAGATAGCCAAAGCTACAAATTTTTTCTGTTTGTCTGCAAT 420		gene		Sequence update by submitter	
DB		24447 GAAAGACAATGGTACCTGAGATAGCCAAAGCTACAAATTTTTTCTGTTTGTCTGCAAT 24388		CDS		On May 24, 2000 this sequence version replaced gi:3192922.	
QY		421 GCTACAAATAGTGATAAATAAGTTTATAGTACTAGAGAATAACGGCGTTGTGACAAAGTCT 480		Location/Qualifiers		1. .2977	
DB		24387 GCTACAAATAGTGATAAATAAGTTTATAGTACTAGAGAATAACGGCGTTGTGACAAAGTCT 24328		/organism="Ehrlichia chaffeensis"		/mol_type="genomic DNA"	
QY		481 CTTAATGTAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540		/strain="Arkansas"		/db_xref="taxon:945"	
DB		24327 CTTAATGTAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 24268		complement (180. .995)		/gene="p28-20"	
QY		541 TGTGCTGGTGTGCTGCAGATTATATAAGTTTTTATAGGTATATCATTTGCTAAGTTTCT 600		complement (180. .995)		/gene="p28-20"	
DB		24267 TGTGCTGGTGTGCTGCAGATTATATAAGTTTTTATAGGTATATCATTTGCTAAGTTTCT 24208		/note="component of 28-kDa outer membrane protein		/codon_start=1	
QY		601 TATCAAAGTTAAGTTTGGTGTCAAACCTCAACCTCTAAATGTTTAATACTATGTTGTTGGTGG 660		/transl_table=11		/product="p28-20 precursor"	
DB		24207 TATCAAAGTTAAGTTTGGTGTCAAACCTCAACCTCTAAATGTTTAATACTATGTTGTTGGTGG 24148		/protein_id="AAC19134.2"		/db_xref="GI:8052525"	
QY		661 GGTATTACCAATAGGTTGTAGTGATAGGCATGAGAGTAGAATAAGCTTACCATCT 720		/translation="MNYKKFVGVVALATLLSFLPDSFSDANVPGRKGFVVGTOYKV		GVNPFNSPASETLPLGLTKSIFALGLDKSSISDHAGFTQAVNPTVASNFAQPGGVIGY	
DB		24147 GGTATTACCAATAGGTTGTAGTGATAGGCATGAGAGTAGAATAAGCTTACCATCT 24088		YVDFRVPFEGAYENFERQWYPEGSESHKFPALSRSTQDNKFIVLNDGVLDKX		LVNPFCDIAGHSIPLDAPYMCAGVADYIKFLGISLSPKFSQVRFQVNPVPSVNVNLF	
QY		721 ACTGCAATTCTGACGTTCTCGAAGTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 780		complement (921. .995)		GGGYHKYVIGNRYERVEIAYHPATLTNVPKTTTSASATLTDYFGVEVGMRETL"	
DB		24087 ACTGCAATTCTGACGTTCTCGAAGTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 24028		complement (183. .920)		/gene="p28-20"	
QY		781 TTTGGTTGGGAGATTGGATTTAGATTGGCGTA 813		/gene="p28-20"		/product="p28-20"	
DB		24027 TTTGGTTGGGAGATTGGATTTAGATTGGCGTA 23995		complement (1178. .2014)		/gene="p28-21"	
RESULT 5				complement (1178. .2014)		/gene="p28-21"	
AF068234/c				/note="component of 28-kDa outer membrane protein		multigene family"	
LOCUS		2977 bp DNA linear BCT 24-MAY-2000		/codon_start=1		/transl_table=11	
DEFINITION		Ehrlichia chaffeensis 28 kDa outer membrane protein gene cluster,		/product="p28-21"		/protein_id="AAF71791.1"	
partial sequence.				/db_xref="GI:8052526"		/translation="MRYRISVIIIMLLLVPCCFSGSLDVSLSNRLLKPVFLGISYK	
AF068234 AF230643				LSAPLFSSFGISGTYRNGVKTDRVVGKSDILLDDAKMDKDNFNFNSEEPYKVDN		NISAPLFSSFGISGTYRNGVKTDRVVGKSDILLDDAKMDKDNFNFNSEEPYKVDN	
AF068234.2 GI:8052524				NLNFGSLFGISGTYRNGVKTDRVVGKSDILLDDAKMDKDNFNFNSEEPYKVDN		NLNFGSLFGISGTYRNGVKTDRVVGKSDILLDDAKMDKDNFNFNSEEPYKVDN	
Ehrlichia chaffeensis				VLKNDGVFYSTILNICYDFADVTNIPVFCVCGIGEDIIKIIPDSIRFKPSPNSKLGIN		VLKNDGVFYSTILNICYDFADVTNIPVFCVCGIGEDIIKIIPDSIRFKPSPNSKLGIN	
				YLMQDMLFFDVFYVYHVRVVGVEYNNIPVQVYSLPNPLNISTAAKLDMVEYFGAIGIKV		YLMQDMLFFDVFYVYHVRVVGVEYNNIPVQVYSLPNPLNISTAAKLDMVEYFGAIGIKV	
						FV"	
				ORIGIN		Query Match	
						70.7%; Score 574.6; DB 1; Length 2977;	

Best Local Similarity 81.7%; Pred. No. 1.4e-94; Matches 664; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy	1	ATGAATATACAAAAGATTGTTGTAGGTGTACGCTGAGTACATTTGTTTTTCTTATCT	60
Db	995	ATGNAATTACAAAATAATTTGTTGAGGTGTCATTTGCTACATGCTTCTTCTGCGCT	936
Qy	61	GATGGTCTTTTCTGTAGCAAAATTTTCTGAGGGAGGAGACATTTATATAGGTAGT	120
Db	935	GATAAATCTTTTCTGATGCAAAAGCTTCTGAAAGGGAGAAAGGATTTTATGTAGGTACT	876
Qy	121	CAGTATAAAGTGTATATCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGCT	180
Db	875	CAATATAAAGTGTGTGTTCTTAATTTTCAGTAATTTTTCAGCTGAAGAAACATTCCTGGA	816
Qy	181	ATTACAAAAAGATTTTGTGCTTGTAGGTCTGTGATAAGTCTGAGATAAATCTCACAGCAAT	240
Db	815	CTCAGAAAAGTATTTTGTGCTTTAGGCTTTGATATAATCTAGCATTAGTGACCATGCAGGT	756
Qy	241	TTTACACGATCATATGACCCCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGATCATTT	300
Db	755	TTTACACAGCGTATAATCTTACATATGCAAGCAATTTTGTGGGTTTGGTGTGTTATT	696
Qy	301	GGATATTCTTANTGACTTTAGGTAGAAATTTGAAGGTTCTTATGAGATTTTGACCT	360
Db	695	GGATATTCTTANTGACTTTAGGTAGAAATTTGAAGGTTCTTATGAGATTTTGACCT	636
Qy	361	GAAAGCAATGTTACCTCTGAGATGAGCAAGCTTACAAATTTTGTGCTCTGCAAT	420
Db	635	GAAAGCAGTGTATCTCTGAGGAGGAGAAAGCATAAGTTTGTCTCTATCTCGTGAG	576
Qy	421	GCTACAAATAGTGATAATAAGTTTATAGTACTAGAGATAACCGCGTTGTGACAGTCT	480
Db	575	TCCACTGTTCAAGATAATAGTTTATAGTGTAGTAAATGATGATGTTTATGATAAATCT	516
Qy	481	CTTAATGTAATGTTTGTATGATATTTGCTAGTGTAGTATCTTCTTAGCACCTTATATG	540
Db	515	CTTAATGTAATGTTTGTATGATATTTGCTAGTGTAGTATCTTCTTAGCACCTTATATG	456
Qy	541	TGCTGCTGTTGTCGAGATTATATAAGTTTGTAGTATATCATTTGCCCTAAGTTTCT	600
Db	455	TGCTGCTGTTGTCGAGATTATATAAGTTTGTAGTATATCATTTGCCCTAAGTTTCT	396
Qy	601	TATCAAGTTAGTTTGGTGTCAACTACCTCTTAAATGTTAATACTATGTTTGTGGGG	660
Db	395	TATCAAGTTAGTTTGGTGTCAACTACCTCTTAAATGTTAATACTATGTTTGTGGGG	336
Qy	661	GTTTATACCAATAGTTTGTAGGTATAGGATAGGATAGGATAGGATAGGATAGGATAGG	720
Db	335	GTTTATACCAATAGTTTGTAGGTATAGGATAGGATAGGATAGGATAGGATAGGATAGG	276
Qy	721	ACTGATTTATCTGAGGTTTCTGAGAACTACTTCACTTCTGCTTCTGCTTCTGCTTCTG	780
Db	275	GCCATTTAAGATGTCCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTG	216
Qy	781	TTTGGTGGAGATTGGATTAGATTTCGGCTA	813
Db	215	TTTGGTGGAGATTGGATTAGATTTCGGCTA	183

RESULT 6  
ECU72291/c  
LOCUS  
DEFINITION  
Ehrlichia chaffeensis strain Arkansas major outer membrane protein  
Omp-1 multigene cluster, complete sequence.  
ACCESSION  
U72291 AF021338  
VERSION  
U72291.2 GI:13511827  
KEYWORDS  
Ehrlichia chaffeensis  
ORGANISM  
Ehrlichia chaffeensis  
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
Anaplasmataceae; Ehrlichia.  
REFERENCE  
1 (bases 14844 to 21136; 21479 to 22234)

AUTHORS  
TITLE  
Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.  
Immunodominant major outer membrane proteins of Ehrlichia chaffeensis are encoded by a polymorphic multigene family  
Infect. Immun. 66 (1), 132-139 (1998)  
JOURNAL  
PUBMED  
9423849  
REFERENCE  
2 (bases 1 to 27190)  
AUTHORS  
TITLE  
Ohashi,N., Rikihisa,Y. and Unver,A.  
Analysis of transcriptionally active gene clusters of major outer membrane protein multigene family in Ehrlichia canis and E. chaffeensis  
Infect. Immun. 69 (4), 2083-2091 (2001)  
JOURNAL  
PUBMED  
11254561  
REFERENCE  
3 (bases 1 to 27190)  
AUTHORS  
TITLE  
Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.  
Direct Submission  
Submitted (24-SEP-1996) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093,  
JOURNAL  
USA  
4 (bases 1 to 27190)  
AUTHORS  
TITLE  
Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.  
Direct Submission  
Submitted (27-AUG-1997) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093,  
JOURNAL  
USA  
5 (bases 1 to 27190)  
AUTHORS  
TITLE  
Ohashi,N., Rikihisa,Y. and Unver,A.  
Direct Submission  
Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093,  
JOURNAL  
USA  
REMARK  
COMMENT  
Sequence update by submitter  
On or before Apr 2, 2001 this sequence version replaced gi:2853584, gi:2853273.  
FEATURES  
source  
1. .27190  
/organism="Ehrlichia chaffeensis"  
/mol\_type="genomic DNA"  
/strain="Arkansas"  
/db\_xref="taxon:945"  
315. .956  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical transcriptional regulator"  
/protein\_id="AAC02808.2"  
/db\_xref="GI:13511828"  
/translation="MSTHAKNKINMDETDDEIKAHDLIAGLIDKMKQKNKFTNKGYN  
DHTSKNFTKGVSTKDKLNKSTTKIRPHPEDECVGKEIKRQIRMRGMSQNLANKLG  
ITFOQVKYKGTNRIVISRLYOLASVLNVEVRDMLKQEDLNKISCDHTVASTSTL  
KNEEFPIBFHDSKIDSKVLMVZAYTCIKNEKVRNIYLVKALSLDKQ"  
1136. .2032  
/genes="omp-1M"  
1136. .2032  
/genes="omp-1M"  
/notes="Omp-1 family member"  
/codon\_start=1  
/transl\_table=11  
/product="major outer membrane protein OMP-1M"  
/protein\_id="AAK28658.1"  
/db\_xref="GI:13511829"  
/translation="MILINMKLVHLDNIMTKFSAIGIVLSVTLFACNVFASPIPID  
FSNESTAGFYASGHYNIQPPRESPIKVKTKDENTKELTFLTKESTDPTFKQS  
BFNDKKGYSPIYNNRYTGFSGAIGYGGURVLEGAFTDFVDKQKIKKDNRYFAL  
CKKDSIESTDNGNHNVMKNEGRVISTFTNACYDMIVSNLSVPSACIGIGGIIIN  
FLGGTNIHTLFAKGLGLGFLSTKPTVIFANGYYVTKNDNSFTNLSVOYPLKEAPKH  
IDPIACFNADNYGGEVGLRFL"  
2498. .3352  
/genes="omp-1N"  
2498. .3352  
/genes="omp-1N"  
/notes="Omp-1 family member"  
/codon\_start=1  
/transl\_table=11  
/product="major outer membrane protein OMP-1N"











Qy	577	GGTATATCATTCGCTAAGTTTTTCTTATCAAGTTTGGTGTCAACTACCCCTTAAAT	636
D5 <td>57102</td> <td>GGAGTTTTCAATTCGCTAAATTTTTCTTATCAGCTTAAATTTGGAGTAACTACCCACTTAAGT</td> <td>57043</td>	57102	GGAGTTTTCAATTCGCTAAATTTTTCTTATCAGCTTAAATTTGGAGTAACTACCCACTTAAGT	57043



Matches 578; Conservative 0; Mismatches 229; Indels 30; Gaps 3;	
Qy	1 ATGAATTACAAAAGATTGTTGTAGGTGTTACCGCTGAGTACATTTGTTTTTCTTAT-- 58
Db	20028 ATGAATTATAAGGAATTGTTTAAAGTGTACATTAAAGTGCATTTGTTTCTTACTA 19969
Qy	59 -CTGATGCTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGACCTTATATAGGT 117
Db	19968 CCTGAACGTGCTATATCTGATATGGAATTTTTCAGAAAATAGAAGTAGTTTATGACAGT 19909
Qy	118 AGTCAGTATAAAGTTGGTATTTCCAAATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCT 177
Db	19908 GTTCAGTATAGGACAGGTATACCAAATTTTCATPAATTTTTCAGCTTCAGAAAACAATCCCT 19849
Qy	178 GGTATTTACAAAAGATTTTTCGCTTAGGCTTTGA-----TAAGTCTGAGATAAATACT 231
Db	19848 GGTTTAACTAAGGGTGTATGGAATTTGGAATCTAAGTAAATCTGATATACAAA 19789
Qy	232 CACAGCAATTTTACACGATCATATGACCCTACTTATCGACGAGTTTTCAGGGTTAGT 291
Db	19788 CGTGCTAATTTTACAGATTAATAATCTTATAGTACCAAGTTCTACAGGGTAGGT 19729
Qy	292 GGTATCATTTGATATTAATGACTTTTAGGGTAGAATTTGAAGTTTCTTATGAGAA 351
Db	19728 GGTATGTTGGGTACTATTTGATATATAAGATGGAATTTGAAACAATCTTATCAAGT 19669
Qy	352 TTTGAACCTGAAAGCAATGGTACCTTGAGAATAGCCAAAGCTACAAATTTTTCGTTTG 411
Db	19668 TTCGGAATTTGAAAGGCAGTGTATCTCTGAGGGAGTCAAAGTCAATAGTTTTCGCTGTA 19609
Qy	412 TCTCGAAATGCTA-----CAAATAGTGATATAAGTTTATAGTA 450
Db	19608 TCTCGTCAAGATTAATGCGAGCTCCTAATACTGACTCAAGTAAATAATAATGTTTGTAGT 19549
Qy	451 CTAGAGAAATACGCGCTTTGTGCAAGTCTTAAATGTAATGTTTCTTATGATATGCT 510
Db	19548 TTGGAATAATAGGGTTTAAGATTAGGACATTCGATGTAATTTTGTATGATGTTGCT 19489
Qy	511 AGTGGTAGTATTCCTTTAGCACCTTATATGTGCTGCTGTTGGTGAGATATATAAAG 570
Db	19488 CACGGAATATTCCTTTAGCACCTTACGTGTGCTGCTGATTTGGTGGGATTAATAA 19429
Qy	571 TTTTGTAGTATATCATTTGCTTAAGTTTCTTATCAAGTTAAGTTTGGTGCAACTACCT 630
Db	19428 TTTATAGGAGTTTCATTCGCTTAAATTTTCTTATCAGCTTAAATTTGAGTAAACTACCA 19369
Qy	631 CTAAATGTTAATACTATGTTTGGTGGGGTTTATACCAAGTTGATAGGTGATAGG 690
Db	19368 CTAAGTATTCGTACTATTTGGTGGGTGGGTATTAATCAATAAGTTATGGGTAGTAAA 19309
Qy	691 CATGAGAGATGAGAAATAGCTTACCATCTACTGCTATTTCTGACGTTCTCTAGAACTACT 750
Db	19308 TATGATAGGTTGAAGTGGTTTATCATCTCTGTTCAAGTTAAATACTGTTCTTAAATGACA 19249
Qy	751 TCAGCTTCTGCTACTTTAAATACATGATATTTTGGTTGGGAGATGGAATTTAGATTT 807
Db	19248 TTTGTTTCTGCTAATTTAGATATTAATTTTGGTTGGTGGGTATTTTGGTTGGGATTAGATTT 19192

RESULT 10  
CR767821\_14/c  
WPCOMMENT

Sequence split into 16 fragments LOCUS CR767821 Accession CR767821

Fragment Name	Begin	End
CR767821_00	1	110000
CR767821_01	100001	210000
CR767821_02	200001	310000
CR767821_03	300001	410000
CR767821_04	400001	510000
CR767821_05	500001	610000
CR767821_06	600001	710000
CR767821_07	700001	810000
CR767821_08	800001	910000

CR767821_09	900001	1010000
CR767821_10	100001	1110000
CR767821_11	1100001	1210000
CR767821_12	1200001	1310000
CR767821_13	1300001	1410000
CR767821_14	1400001	1510000
CR767821_15	1500001	1516355
Continuation (15 of 16) of CR767821 from base 1400001 (CR767821 Ehrlichia ruminantium str		
Query Match 46.8%; Score 380.6; DB 1; Length 110000;		
Best Local Similarity 69.1%; Pred. No. 6.1e-60;		
Matches 578; Conservative 0; Mismatches 229; Indels 30; Gaps 3;		
Qy	1 ATGAATTACAAAAGATTGTTGTAGGTGTTACCGCTGAGTACATTTGTTTTTCTTAT-- 58	
Db	93551 ATGAATTATAAGGAATTTGTTTTAGGTGTACATTAAAGTGCATTTGTTTCTTACTTA 93492	
Qy	59 -CTGATGCTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGAGACTTTATATAGGT 117	
Db	93491 CTTGAACGTGCTATATCTGATATGGAATTTTTCAGAAAATAGAAGTAGTTTATGACAGT 93432	
Qy	118 AGTCAGTATAAAGTTGGTATTTCCAAATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCT 177	
Db	93431 GTTCAGTATAGGACAGGTATACCAAATTTTGTATAATTTTTCAGCTTCAGAAAACAATCCCT 93372	
Qy	178 GGTATTTACAAAAGATTTTTCGCTTAGGCTTTGA-----TAAGTCTGAGATAAATACT 231	
Db	93371 GGTTTAACTAAGGGTGTATTTATGGAATTTGGAATCTAAGTAAATCTGATATACAAA 93312	
Qy	232 CACAGCAATTTTACACGATCATATGACCCTACTTATGCAAGCAGTTTTCAGGGTTTAGT 291	
Db	93311 CGTGCTAATTTTACAGATTAATAATCTTATAGTACCAAGTTCTACAGGGATAGGT 93252	
Qy	292 GGTATCATTTGATATTAATGACTTTTAGGGTAGAATTTTCAAGTCTTATGAGAAAT 351	
Db	93251 GGTATGTTGGGTACTATTTTGTATAATAAGATGGAATTTTGAACAATCTTATCAAGT 93192	
Qy	352 TTTGAACCTGAAAGCAATGGTACCTTGAGAATAGCCAAAGCTACAAAATTTTTCGCTTGG 411	
Db	93191 TTTGGAATTTGAAAGGCAGTGGTATCTCTGAAGGGAGTCAAAGTCATAAGTTTTCGCTGTA 93132	
Qy	412 TCTCGAAATGCTA-----CAAATAGTGATATAAGTTTATAGTA 450	
Db	93131 TCTCGTCAAGATTAATGCGAGCTCCTAATACTGACTCAAGTAAATAATAATGATTTGTAGT 93072	
Qy	451 CTAGAGAAATAACGCGCTTGTGCAAGTCTCTTAATGTAATGTTTGTATGATATGCT 510	
Db	93071 TTGGAATAATAAGGGTTTAAGATTAGGACATTTGCATGTAATAATTTTGTATGATGCT 93012	
Qy	511 AGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGTTGGTGGCAGATTAATAAAG 570	
Db	93011 CACGGAATATTCCTTTAGCACCTTACGTGTGCTGCTGATTTGGTGGGATTAATAA 92952	
Qy	571 TTTTGTAGTATATCATTTGCTTAAGTTTCTTATCAAGTTTAAAGTTTGGTCACTACCCT 630	
Db	92951 TTTATAGGAGTTTCATTTGCTTAAATTTTCTTATCAGCTTAAATTTTGGAGTAAACTACCCA 92892	
Qy	631 CTAAATGTTAATACTATGTTTGGTGGGGTTATTACCATAAGGTGTAGGTGATAGG 690	
Db	92891 CTAAGTATTCGTACTATGCTATTTGGTGGTGGGTATTAATCAATAAGTTATGGGTAGTAAA 92832	
Qy	691 CATGAGAGATGAGAAATAGCTTACCATCTACTGCAATTAATCTGACCTTCTAGAACTACT 750	
Db	92831 TATGATAGGTTGAAAGTGGTTTATCATCTGTTTCAAGTTAAATCTGTTCTTAAATGACA 92772	
Qy	751 TCAGCTTCTGCTACTTTAAATACATGATATTTTGGTTGGGAGATTTGGAATTTAGATTT 807	
Db	92771 TTTGTTTCTGCTAATTTAGATATTAATTTTGGTTGTGTAGGTGGGATTAGATTT 92715	
RESULT 11		
CR925678_14/c		
WPCOMMENT		

Sequence split into 16 fragments LOCUS CR925678 Accession CR925678

Fragment Name	Begin	End
CR925678_00	1	110000
CR925678_01	100001	210000
CR925678_02	200001	310000
CR925678_03	300001	410000
CR925678_04	400001	510000
CR925678_05	500001	610000
CR925678_06	600001	710000
CR925678_07	700001	810000
CR925678_08	800001	910000
CR925678_09	900001	1010000
CR925678_10	1000001	1110000
CR925678_11	1100001	1210000
CR925678_12	1200001	1310000
CR925678_13	1300001	1410000
CR925678_14	1400001	1510000
CR925678_15	1500001	1512977

Continuation (15 of 16) of CR925678 from base 1400001 (CR925678 Ehrlichia ruminantium st

Query Match	46.8%;	Score 380.6;	DB 1;	Length 110000;
Best Local Similarity	69.1%;	Pred. No. 6.1e-60;		
Matches 578;	Conservative 0;	Mismatches 229;	Indels 30;	Gaps 3;

QY	1	ATGAATTACAAAAGATTTGTTGAGGTGTACGCTCAGTACATTTGTTTCTTTCTTAT-- 58
Db	71320	ATGAATTATAGAAATTTGTTTGGTGTTACATTAAAGTCATTTGTTCTTTACTA 71261
QY	59	-CTGATGTTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGAGACTTTATATAGGT 117
Db	71260	CCTGAAGCTGTATATCTGATATGGATGTTTCAGAAAATAGAAAGTAGGTTTTATGCAGT 71201
QY	118	AGTCAGTATAAAGTTGGTATTCCTCAATTTTAGTAATTTTTCAGCTGAAGAAACAATCTCT 177
Db	71200	GTTTCAGTATAGGACAGGTATACCAAAATTTTGATAATTTTTCAGCTTCAGAAAACAATCCCT 71141
QY	178	GGTATTACAAAAAGATTTTGGTTAGGTCTTGA-----TAAGTCTGAGATAAATACT 231
Db	71140	GGTTAACTAAGGGTGTATTGAGATTAGATTGGATCTAAGTAATCTGATATACAAA 71081
QY	232	CACAGCAATTTTACAGCATATAGACCCCTACTTATGCAAGCAGTTTTCAGGGTTTACT 291
Db	71080	CGTGCTAAATTTTCAAGATTATATAATCTCTACTTATAGTACCAGTTCTACAGGATAGT 71021
QY	292	GGTATCATTTGGATTTATGTTATGATCTTTAGGGTAGAATTTGAAAGTTCTTATGAGAT 351
Db	71020	GGTATGTTGGGTACTATTTTGATAATATAGAAGTGGAAATTTGAAACATCTTATTTCAAGT 70961
QY	352	TTTGAACCTGAAAGACAATGGTACCTCGTGAAGATAGCAAGCTACAAAATTTTGTCTTG 411
Db	70960	TTTCGGAATTGAAGGCAGTGATCTCTGAAGGAGTCAAAAGTCATGAAGTTTGTGCTGTA 70901
QY	412	TCTCGAAATGCTA-----CAAATAGTGATATAAGTTTATAGTA 450
Db	70900	TCTCGTCAAGATAATGTCAGCTCCTAATACGACTCAAGTAAATAATGATTTGTAGTG 70841
QY	451	CTAGAGAAATACCGCGTTGTTGACAAGTCTCTTAATGTTAAATGTTGTTATGATATGCT 510
Db	70840	TTTGAAGAAATTAATCGGGTAAAGATTAGGACATTCGATGTAATTTTGTATGATGTGCT 70781
QY	511	AGTGGTAGTATTCCTTTAGCACCTTATATGTTGCTGGTGTGCTGCAGATTATATAAAG 570
Db	70780	CACGGAATATTCCTTTAGCACCTTACGTTGCTGCTGTTATGTTGGGATTAATGTA 70721
QY	571	TTTTTAGGTATATCATTTGCTAAAGTTTCTTATCAAGTTAAGTTTGGTGCACACTACCT 630
Db	70720	TTTTTAGGAGTTTCATTGCTCAAAATTTCTTATCAGCTTAAATTTTGGAGTAAACTACCCA 70661
QY	631	CTAAATGTTAATACTATGTTTGGTGGGGTTATACCATAGTTTGTAGTGATAGG 690
Db	70660	CTAAGTATTCGTACTATGCTATTTTGGTGGGGGTATATCATCAAGTTATGGGTAGTAA 70601
QY	691	CATGAGAGAGTGAATAAGCTTACCCTACTGTCATTTATCTGACGTTCTCTGAAGACTACT 750

Db	70600	TATGATAGGTGCAAAAGTGGTTATCATCTGTTCAAGTTAAATCTGTTCTTAAATGACA 70541
QY	751	TCAGCTTCTGCTACTTTAAATCTGATTTATTTGTTGGGAGATTGGATTTAGATTT 807
Db	70540	TTTGTTTCTGCTAAATTTAGATATTGATTTATTTGTTGTTGAGGTTGGGATTAGATTT 70484
RESULT 12		
AF319940/c		
LOCUS	19449 bp	DNA linear BCT 05-JUL-2005
DEFINITION	Ehrlichia ruminantium isolate Senegal map1 gene cluster, partial sequence.	
ACCESSION	AF319940	
VERSION	AF319940.4	GI:66957871
KEYWORDS	Ehrlichia ruminantium (heartwater rickettsia)	
SOURCE	Ehrlichia ruminantium	
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Ehrlichia.	
REFERENCE	1 (bases 1 to 19449)	
AUTHORS	Bekker,C.P., Bell-Sakyi,L., Paxton,E.A., Martinez,D., Bensaïd,A. and Jongejan,F.	
TITLE	Transcriptional analysis of the major antigenic protein 1 multigene family of Cowdria ruminantium	
JOURNAL	Gene 285 (1-2), 193-201 (2002)	
PUBMED	12039046	
REFERENCE	2 (bases 1 to 19449)	
AUTHORS	Bekker,C.P., Postigo,M., Taoufik,A., Bell-Sakyi,L., Ferraz,C., Martinez,D. and Jongejan,F.	
TITLE	Transcription Analysis of the Major Antigenic Protein 1 Multigene Family of Three In Vitro-Cultured Ehrlichia ruminantium Isolates	
JOURNAL	J. Bacteriol. 187 (14), 4782-4791 (2005)	
PUBMED	15995193	
REFERENCE	3 (bases 1 to 19449)	
AUTHORS	Bekker,C.P.J., Paxton,E.A., Bell-Sakyi,L., Martinez,D., Bensaïd,A. and Jongejan,F.	
TITLE	Direct Submission	
JOURNAL	Submitted (07-NOV-2000) Bacteriology, Faculty of Veterinary Medicine, Yalelaan 1, Utrecht 3584CL, The Netherlands	
REFERENCE	4 (bases 1 to 19449)	
AUTHORS	Bekker,C.P.J., Taoufik,A. and Jongejan,F.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-NOV-2004) Virology, Utrecht University, Yalelaan 1, Utrecht 3584CL, The Netherlands	
REMARK	Sequence update by submitter	
COMMENT	On Jun 6, 2005 this sequence version replaced gi:55620638.	
FEATURES	Location/Qualifiers	
source	1..19449	
	/organism="Ehrlichia ruminantium"	
	/mol_type="genomic DNA"	
	/isolate="Senegal"	
	/db_xref="taxon:779"	
gene	600..1523	
	/gene="map1-14"	
CDS	600..1523	
	/gene="map1-14"	
	/note="member of multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="MAP1-14"	
	/protein_id="AAV59045.1"	
	/db_xref="GI:66957874"	
	/translation="MDNVVVFSCGLGFALSLVTLVMAHNAPSSPLPVDFSESENVGFP YTSQYSLEVFPKFSVAVKYKHEKODKELTIFSLKBENTELKNDKGFKKGYNPVYN RNYTGFSAIGYSGGLRIELEGATKFDQKQYFQDNVRYFALSKDEEISGGPDQ PTPVPVPVPAQPSKPTTGYNVVTAKEGLSIISLTNACYDVIIGNSOLIPSVCI IGIGGINTFLGVINIKTIYKAKVGMGLLSPKTIIFINGVYVVKVPNDSFKNVSIXYQ HELEKDPKHBPFIIPFNDSYDYGGEVGLRFL"	
	2400..3284	
gene	/gene="map1-13"	
	2400..3284	
CDS	/gene="map1-13"	



```
QY 103 GGACTTTATATAGGTAGTCAGTATAAAGTTGGTATTCCTCCAAATTTTAGTAATTTTCAGCT 162
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19367 AGGTTTATGCGAGTGTTCATATAGCAGGTATACCAATTTTGTATATTTTCAGCT 19308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 163 GAAGAAACAATTCCTCGTATTACAAAAGAGATTTTTCGGTTAGGCTTCGA-----TAAG 216
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19307 TCAGAAACAATCCCTGGTTTAACTAAGGGGTGTTTATGGATTAGATTGGATCTAAGATAA 19248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 217 TCTGAGATAAATCTCACAGCAATTTTACAGCATATATGACCCCTACTTATGCAAGCAGT 276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19247 TCTGGTATTACAAAACGTGCTAAATTTTACAAGATTGTATTAATCCTACTTTATAGTACTAGT 19188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 TTTGCAAGGTTTATGCGTATCATTTGGATATATGTTTAATGCACTTTAGGGTAGAATTTGAA 336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19187 TCTACAGGATAGGTGTTATGTTGGGTACTATTTTGTATATATAGAAATGAGTTTGA 19128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 GGTTCCTATAGAAATTTTGAACCTGAAAGCAATGGTACCCCTGAGAAATGCCAAGCTAC 396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19127 ACATCTATTCAAATTTCCGAGTTGAAAGGCGAGTGGTATCCTGAAGGGAGTCAAAGTCAT 19068
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 397 AAATTTTTCCTTGTCTCGAAATGC-----TACAAATAGTANTATAG 441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19067 AAGTTCTGTGCTGTATCTCGTAATGCAACTCCTAGTACTGATTCAGATTAATAATGAT 19008
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 TTTATAGTACTAGAGAAATACGGCGTTGTTGACAAAGTCTCTTAATGTAAATGTTCTTAT 501
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19007 TTTGTAGTCTAGAAATAATAGGGTTAAGGTTAGGACATTAAGATGTAATTTTGTAT 18948
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 502 GATATTGCTAGTGGTATATTCCTTTAGCACCTTATATGTGTCGTGTTGTCAGAT 561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18947 GATGTTGCTCACGGAATATTCCTTTAGCACCTTATGTTGTCGTGTTGTCGGGAT 18888
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 TATATAAGTTTATAGTATATCATTTGCCCTAAGTTTCTTATCAAGTTAAGTTGTCGTC 621
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18887 TATGTAAATTTATAGGAGTTTCATTCGCCCTAAATTTTCTTATCAGCTTAAATTTGGAGTA 18828
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 622 AACTACCCTCTAAATGTTTAATCTATGTTGTTGGTGGGGTTATTACCAATAAGTTGTA 681
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18827 AACTATCCACTAGTATTCGTACTATGCTATNTTGGTGGGGTTATTCATAAAGTTATA 18768
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 682 GGTGATAGGATGAGAGAGTAAAGTATGCTTACCATCTACTGCTATTCAGTTCCT 741
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18767 GGTAGTAAATATGATAGGGTAAAGTAGTTTATCATCTGTTCAAGTTAAGTACTGTTCT 18708
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 742 AGAACTACTTCAGCTTCTGCTACTTTAAATCTAGTATTTTGGTTGGGAGATTGATTT 801
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18707 ANAATGACATTTGTTCTGCTAAATTTAGATATTTGATTTTGGTTGTGAAGTTGGGATT 18648
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 802 AGATTT 807
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18647 AGATTT 18642
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 13
AR303095 LOCUS 840 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 23 from patent US 6544517.
ACCESSION AR303095
VERSION AR303095.1 GI:31691735
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 840)
AUTHORS Rikihisa,Y. and Ohashi,N.
TITLE Outer membrane protein of Ehrlichia canis and Ehrlichia chaffeensis
JOURNAL Patent: US 6544517-A 23 08-APR-2003;
The Ohio State University Research Foundation; Columbus, OH
FEATURES
source Location/Qualifiers
1..840
/mol_type="genomic DNA"
/organism="unknown"
```

```
Query Match 20.0%; Score 162.2; DB 6; Length 840;
Best Local Similarity 53.8%; Pred. No. 1.3e-19;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTACGCTGAGTACATTTGTTTCTTATCTGATGGTGTCTTTTCTGATGCA 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 34 GTACTTGCATCTCTATTATCATTTCTTATCTATTGAATCCTTTTCAGCTATAAATCATAT 93
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 82 AATTTTCTGAAGGAGGAGAGAGACTTTATATAGTAGTCAGTATAAAGTTGGTATTTCCC 141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 94 CATACAGAAATAAACACTAGTGTATATATATACAGGCGAGTATAGACCAGAGTATCC 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 142 AATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGCTATTACAAAAAGATTTTGGG 201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 154 CATTTTGTAGCAATTTCTCAGTAAAGAACTAATGTTGTATACAACTAGTAGATAT 213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 202 TTAGGTCTTGATAAGTCT---GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 214 AAAAAAGTGGCTCTTCTATCGATCCTAACACTTATTCAAAGTCCATATACT 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 259 CCTACTATGCAAGCAGGTTTTGCAGGGTTTAGTGTATCATTTGGATATTTATGTTAATGAC 318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 274 GTTACATTTCAAGATAATGCTGTAGTTTTCAGTGGAGCAATTTGGATATTTCTTACCCGAA 333
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 319 TTT---AGGGTGAATTTGAAAGTTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTAC 375
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTTGATGTCAAAGATCCTAAAGAC 393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 CCTGAGAATAGCCAAAGCTACAAATTTTGTCTCGAAATGCTACAAATAGTATGAT 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 394 TACTCAGCAAAAGATGCTTTTAGGTTTTCCTAGCACGTAATACTGCTACTACTGTT 453
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 436 AAT-----AAGTTTATAGTACTAGAGAAATAACGGGTTGTTGACAAGTCTCTTAAT 486
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 454 CCTGATGCTCAAAAAATATACAGTTATGAAGAAATAATGCTTATCTGTCATCAATCATG 513
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 GTAAATGTTTGTATGATATGCTAGTGTAGTATTTCTTTTAGCACCTTATATATGTTGCT 546
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 514 ATCAATGTTGTTATGATCTATCTTTTAATTAATTTAGTGTATCATCCTTATATATGCA 573
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 GGTGTTGTTGCAAGATATATAAAGTTTATAGGTATATCATTCGCTAAGTTTCTTATCAA 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 574 GGTATGTTGTTGAGATTTTCATTGAATTTTGTATCTTTGCAACATTTAAACTTCTTATCAA 633
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 GTTAAGTTTGGTGTCAATACCCCTCTAAATGTTTAATATCTATGTTGTTGGTGGGGTTAT 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 634 GGAATAACTAGGTATTTAGTTTATTTACTTCTTCTTCTAAGATTAATGTTGTTGCTGGGTAC 693
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 667 TACCATAAGGTTGTAGGTGATAGGCATGAGAGAGTAGAAATAGCTTACCATCCTACTGCA 726
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 694 TATCATAGATTATAGGGAATAAATTTAAAAATTT---AATGTTAAACCATGTTGTACA 750
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 727 TTATCTGAGTTCTTAGAACTACTTTCAGCTTCTGCTACTTTTAAATACGATTAATTTTGT 786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 751 CTTGATGAATTTCTTAAGCACTTCTGAGTAGTACACTTAATGTTGCTTATTTTGT 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 787 TGGGAGATTGGATTTAGATTTGC 809
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 811 GGTGAAGCTGGAGTAAAGTTTAC 833
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AR655791 LOCUS 840 bp DNA linear PAT 13-JUN-2005
DEFINITION Sequence 23 from patent US 6893640.
ACCESSION AR655791
VERSION AR655791.1 GI:67588223
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 840)
Unclassified.
```









Db 7651 AAAAAAGTGGCTCTTCTATCGATCCATAACACTTATTCAAACCTTTCAAGGTCCATATACT 7710  
Qy 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGGTATCAATGCGATATATGTTAATGAC 318  
Db 7711 GTTACATTTCAAGATAATGCTGCTAGTTCAGTGGAGCAATTGGATATCTTACCCGAA 7770  
Qy 319 TTT---AGGGTAGAATTTGAGGTTCTTATGAGAAATTTGAACTTGAAGCAATGGTAC 375  
Db 7771 AGCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTGATGTCAAAGATCCTAAAGAC 7830  
Qy 376 CCTGAGAAATAGCCAAAGCTACAAATTTTTCCTTTGCTCGAAATGCTACAAATAGTAT 435  
Db 7831 TACTCAGCAAAAGATGCTTTTAGGTTTTCCTCTAGCAGTAATAGTCTACTACTGTT 7890  
Qy 436 AAT-----AAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCTCTTAAT 486  
Db 7891 CCTGATGCTCAAAATAATACAGTTATGAAGAATAATGGCTTATCTGTTGCATCAATCATG 7950  
Qy 487 GTAATGTTGTTTATGATAATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCT 546  
Db 7951 ATCAATGGTGTATGATCTATCTTTTAATAAATTTAGTCGTATCACCTTATATATGCA 8010  
Qy 547 GGTGTGGTGCAGATTATATAAAGTTTTCAGGTATATCATTCCTTAAGTTTTCCTTACAA 606  
Db 8011 GGTATGGTGAAGATTTTCATTTGAATTTTTCATCTTGCACATTAACCTTGCTTATCAA 8070  
Qy 607 GTTAAGTTTGGTGTCAACTACCCCTCTAAATGTTAATACTATGTTGTTGGTGGGGTTAT 666  
Db 8071 GGAAAACTAGGTATTAGTTATTACTTCTTCCCTAAGATTAATGTTATTTGCTGGTGGGTAC 8130  
Qy 667 TACCATAAGTTGTAGGTGATAGGCTAGAGAGATAGAAATAGCTTACCATCCTACTGCA 726  
Db 8131 TATCATAGATTATAGGGAAATAAATTTTAAAAATTT---AAATGTTAAACCATGTTGTTACA 8187  
Qy 727 TTATCTGACGTTCCCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATACTGATTATTTGGT 786  
Db 8188 CTTGATGAATTTCCCTAAGCAACTTCTGCAGTAGCTACACTTAATGTTGCTTATTTTGGT 8247  
Qy 787 TGGGAGATTGGATTAGATTTC 809  
Db 8248 GGTGAAGCTGGAGTAAAGTTTAC 8270

Search completed: December 28, 2005, 12:12:07  
Job time : 4522 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 28, 2005, 10:46:11 ; Search time 620 Seconds  
(without alignments)  
8739.347 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaagaattgtt.....ttggattgattggcgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_21.\*
- 1: Geneseq1980s.\*
  - 2: Geneseq1990s.\*
  - 3: Geneseq2000s.\*
  - 4: Geneseq2001as.\*
  - 5: Geneseq2001bs.\*
  - 6: Geneseq2002as.\*
  - 7: Geneseq2002bs.\*
  - 8: Geneseq2003as.\*
  - 9: Geneseq2003bs.\*
  - 10: Geneseq2003cs.\*
  - 11: Geneseq2003ds.\*
  - 12: Geneseq2004as.\*
  - 13: Geneseq2004bs.\*
  - 14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813	100.0	813	6	ABK68878 DNA encod
2	162.2	20.0	840	2	AAX34754 DNA encod
3	162.2	20.0	840	6	ABS63287 DNA encod
4	162.2	20.0	840	9	ADA09756 E. chaffe
5	162.2	20.0	840	14	ADM04249 Ehrlichia
6	150.8	18.5	852	2	AAX34755 DNA encod
7	150.8	18.5	852	6	ABS63288 DNA encod
8	150.8	18.5	852	9	ADA09758 E. chaffe
9	150.8	18.5	852	14	ADM04251 Ehrlichia
10	147.2	18.1	840	6	ABK68876 DNA encod
11	147.2	18.1	843	6	ABS63299 DNA encod
12	147.2	18.1	843	9	ADA09780 DNA encod
13	147.2	18.1	843	14	ADM04273 Ehrlichia
14	146.6	18.0	726	6	ABS63306 DNA encod
15	146.6	18.0	726	9	ADA09794 E. canis
16	146.6	18.0	726	14	ADM04287 Ehrlichia
17	146.4	18.0	849	6	ABT11835 E chaffe
18	146.4	18.0	849	14	ABE95350 Ehrlichia
19	135	16.6	828	2	AAX34756 DNA encod

20	135	16.6	828	6	ABS63289 DNA encod
21	135	16.6	828	9	ADA09760 Ehrlichia
22	135	16.6	828	14	ADM04253 Ehrlichia
23	133.8	16.5	840	6	ABS63305 DNA encod
24	133.8	16.5	840	9	ADA09792 E. canis
25	133.8	16.5	840	14	ADM04285 Ehrlichia
26	129.6	15.9	843	2	AAX34745 DNA encod
27	129.6	15.9	843	6	ABS63278 DNA encod
28	129.6	15.9	843	9	ADA09738 E. chaffe
29	129.6	15.9	843	14	ADM04231 Ehrlichia
30	126.4	15.5	837	2	AAX34747 Ehrlichia
31	126.4	15.5	837	6	ABS63280 DNA encod
32	126.4	15.5	837	9	ADA09742 E. chaffe
33	126.4	15.5	837	14	ADM04235 Ehrlichia
34	126.4	15.5	837	14	ADM04257 Ehrlichia
35	124.8	15.4	4683	3	AAC68716 Ehrlichia
36	124.8	15.4	4683	3	AAV07179 Ehrlichia
37	123.2	15.2	4683	2	AAV07179 Ehrlichia
38	122.4	15.1	867	2	AAX34759 DNA encod
39	122.4	15.1	867	6	ABS63291 DNA encod
40	122.4	15.1	867	9	ADA09764 E. canis
41	122.4	15.1	867	14	ADM04257 Ehrlichia
42	121.8	15.0	843	2	AAX34748 DNA encod
43	121.8	15.0	843	3	AAC68705 Ehrlichia
44	121.8	15.0	843	6	ABS63281 DNA encod
45	121.8	15.0	843	9	ADA09744 E. chaffe

ALIGNMENTS

RESULT 1  
ABK68878  
ID ABK68878 standard; DNA; 813 BP.  
XX  
AC ABK68878;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE DNA encoding Ehrlichia canis p28-9.  
XX  
KW Ehrlichia canis infection; vaccine; serodiagnostic; gene; p28; ss;  
KW antibacterial.  
XX  
OS Ehrlichia canis.  
XX  
PN WO200222782-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 12-SEP-2001; 2001WO-US028759.  
XX  
PR 12-SEP-2000; 2000US-00660587.  
XX  
PA (RERE-) RES DEV FOUND.  
XX  
PI Walker DH, Yu X, McBride JW;  
XX  
DR WPI; 2002-351882/38.  
XX  
P-PSDB; AAU96118.  
XX  
New recombinant homologous 28 kilodalton immunodominant protein from Ehrlichia canis, useful for treating Ehrlichia canis infections.  
XX  
Claim 5; Fig 16; 106pp; English.

XX The invention relates to a recombinant homologous 28 kDa immunodominant protein, p28 (I) of Ehrlichia canis. (I), a 28-kDa antigen preferably dispersed in a pharmaceutically acceptable carrier, is useful for inhibiting E. canis infection in a subject. (I) is useful in the development of vaccines and serodiagnostics that are particularly effective for disease prevention and serodiagnosis. ABK68878 represent the 28-kDa antigen coding sequences and PCR primers of the

```
CC invention
XX
SQ Sequence 813 BP; 232 A; 106 C; 171 G; 304 T; 0 U; 0 Other;

Query Match      100.0%; Score 813; DB 6; Length 813;
Best Local Similarity 100.0%; Pred. No. 4.1e-182;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATACAAAGATTGTTGAGGTGTACGCTGAGTACATTTGTTTTTCTTATCT 60
   |||||
DB 1 ATGAATACAAAGATTGTTGAGGTGTACGCTGAGTACATTTGTTTTTCTTATCT 60

QY 61 GATGGTGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGAGACTTTATATAGTAGT 120
   |||||
DB 61 GATGGTGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGAGACTTTATATAGTAGT 120

QY 121 CAGTATAAAGTTGGTATTCCTCAATTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGT 180
   |||||
DB 121 CAGTATAAAGTTGGTATTCCTCAATTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGT 180

QY 181 ATTACAAAAGATTTTGGGTTAGTCTTGATAAGTCTGAGATAAATCTACAGCAAT 240
   |||||
DB 181 ATTACAAAAGATTTTGGGTTAGTCTTGATAAGTCTGAGATAAATCTACAGCAAT 240

QY 241 TTTACAGATCATATGACCCCTACTTATGCAAGCAGTTTTTCAGGGTTTAGTGTATCAT 300
   |||||
DB 241 TTTACAGATCATATGACCCCTACTTATGCAAGCAGTTTTTCAGGGTTTAGTGTATCAT 300

QY 301 GGTATTTATGTTAATGATTTAGGGTAGAATTTGAAGGTTCTTATGAGAATTTGAACCT 360
   |||||
DB 301 GGTATTTATGTTAATGATTTAGGGTAGAATTTGAAGGTTCTTATGAGAATTTGAACCT 360

QY 361 GAAAGCAATGGTACCTGAGATAGCCAAAGCTACAAATTTTTCGTTTGTCTGCAAT 420
   |||||
DB 361 GAAAGCAATGGTACCTGAGATAGCCAAAGCTACAAATTTTTCGTTTGTCTGCAAT 420

QY 421 GCTCAAAATAGTATTAATAGTTTATAGTACTAGAGATAACGGCGTTGTTGCAAGTCT 480
   |||||
DB 421 GCTCAAAATAGTATTAATAGTTTATAGTACTAGAGATAACGGCGTTGTTGCAAGTCT 480

QY 481 CTTAATGTAATGTTTGTATGATATTTGCTAGTGTAGTATTCCTTTAGCACCTTATATG 540
   |||||
DB 481 CTTAATGTAATGTTTGTATGATATTTGCTAGTGTAGTATTCCTTTAGCACCTTATATG 540

QY 541 TGTGCTGGTGTGGTGCAGATTATATAAGTTTTTTTAGGTATATCATTCGCTAAGTTTTCT 500
   |||||
DB 541 TGTGCTGGTGTGGTGCAGATTATATAAGTTTTTTTAGGTATATCATTCGCTAAGTTTTCT 500

QY 601 TATCAAGTTAAGTTTGGTGTCAACTACCTCTCAATAGTTTAAATAGTTTGTGGTGG 560
   |||||
DB 601 TATCAAGTTAAGTTTGGTGTCAACTACCTCTCAATAGTTTAAATAGTTTGTGGTGG 560

QY 661 GGTATTACCAATAGGTTGTAGTGATAGGCATGAGAGATAGAAAATAGCTTACCATCT 720
   |||||
DB 661 GGTATTACCAATAGGTTGTAGTGATAGGCATGAGAGATAGAAAATAGCTTACCATCT 720

QY 721 ACTGCATTATCTGACGTTCTGACAACTTCTGAGTCTTCTGCTACTTTAAATAGTATT 780
   |||||
DB 721 ACTGCATTATCTGACGTTCTGACAACTTCTGAGTCTTCTGCTACTTTAAATAGTATT 780

QY 781 TTTGGTTGGGAGATTGGATTAGATTGCGCTA 813
   |||||
DB 781 TTTGGTTGGGAGATTGGATTAGATTGCGCTA 813
```

```
RESULT 2
AAX34754
ID AAX34754 standard; DNA; 840 BP.
XX
AC AAX34754;
XX
DT 27-AUG-2003 (revised)
DT 05-JUL-1999 (first entry)
```

```
XX DNA encoding OMP-1V protein.
DE
XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
KW detection; dog; ss.
XX Ehrlichia chaffeensis.
OS
XX WO9913720-A1.
XX 25-MAR-1999.
XX 18-SEP-1998; 98WO-US019600.
XX 19-SEP-1997; 97US-0059353P.
XX (OHIS ) UNIV OHIO STATE.
XX Rikihisa Y, Ohashi N;
XX WPI; 1999-254290/21.
XX P-PSDB; AAY06954.
XX Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
   canis.
XX Disclosure; Fig 14A; 55pp; English.
XX The invention provides isolated outer membrane proteins (OMP) from
   Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
   of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in
   CC AAY06943-958. The E. canis proteins form part of the P30 family and
   consist of proteins shown in AAY06959-970. The proteins and genes are
   used to detect E. chaffeensis in patients and E. canis in dogs. (Updated
   on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 840 BP; 279 A; 128 C; 136 G; 297 T; 0 U; 0 Other;

Query Match      20.0%; Score 162.2; DB 2; Length 840;
Best Local Similarity 53.8%; Pred. No. 2.1e-28;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTACGCTCAGTACATTTCTTTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
   |||||
DB 34 GTACTTGATCTCTATTATCAITCTTATCTATGATCTTTTCAGCTATAAATCATAT 93

QY 82 AATTTTTCTGAAGGAGGAGGAGACTTTATATAGTAGTCAATATAAAGTTGGTATCCC 141
   |||||
DB 94 CATACAGGAATAACACTAGTGGTATATATATACAGGCGAGTATAGACCAGGATATCC 153

QY 142 AATTTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAGATTTTGG 201
   |||||
DB 154 CATTTTAGCAATTTCTCAGTAAAGAAACTAATGTTGATACAATAACAACCTAGTAGATAT 213

QY 202 TTAGGTCTTGATAAGTCT---GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258
   |||||
DB 214 AAAAAAGTGGCTCTCTATCGATCCTAACACTTATTCAACTTTCAAGTCCATATCT 273

QY 259 CCTACTTATGCAAGCAGTTTTTGCAGGGTTTGTAGTGTATCATTTGGATATTATGTAATGAC 318
   |||||
DB 274 GTTACATTTCAAGATAATGCTGCTAGTTTTCAGTGGAGCAATTTGGATATTTCTACCCGAA 333

QY 319 TTT---AGGTAGATTTGAAAGTTCTTATGAGATTTTGAACCTTGAACGAAAGAGTGGTAC 375
   |||||
DB 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTTGATGTCAAAAGATCCTAAAGAC 393

QY 376 CCTGAGAATAGCAAAAGCTACAAAATTTTGTCTTGTCTCGAAATGCTACAAATAGTGAT 435
   |||||
DB 394 TACTCAGCAAAAGATGCTTTTAGGTTTGTCTAGCACGTAATACGCTACTACTGTT 453

QY 436 AAT-----AAGTTTATAGTACTAGAGATAAACGGGTTTGTGCAAGTCTCTTAAT 486
   |||||
DB 454 CCTGATGCTCAAAAAATATACAGTTATGAAGATAATGCTTATCTGTTGCATCAATCATG 513
```



KW	monocytic ehrlichiosis; Rocky Mountain spotted fever;	
KW	canine ehrlichiosis.	
XX	Ehrlichia chaffeensis.	
OS	US6544517-B1.	
PN		
XX		
PD	08-APR-2003.	
XX		
PF	19-MAY-1999; 99US-00314701.	
XX		
PR	18-SEP-1998; 98US-0100843P.	
XX		
PA	(OHIS ) UNIV OHIO STATE RES FOUND.	
XX		
PI	Rikihisa Y, Ohashi N;	
XX		
DR	WPI; 2003-553952/52.	
DR	P-PSDB; ADA09757.	
XX		
PT	New isolated polynucleotide encoding outer membrane protein P30 of	
PT	Ehrlichia canis or its variant or fragment, useful for producing	
PT	Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for	
PT	diagnosing and treating ehrlichiosis.	
XX		
PS	Disclosure; Fig 14; 105pp; English.	
XX		
CC	The invention relates to an isolated polynucleotide encoding a variant of	
CC	the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer	
CC	membrane protein of E. canis , or an antigenic fragment of the E. canis	
CC	P30 protein, or comprising a sequence which is the complement of	
CC	nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.	
CC	canis and E. chaffeensis outer membrane proteins and their encoding	
CC	nucleic acids. The polynucleotides are useful for producing E. canis or	
CC	E. chaffeensis outer membrane protein, for designing hybridisation probes	
CC	for isolating and identifying cDNA and genomic clones encoding the OMP)	
CC	or its allelic forms, for designing primers for PCR. The polypeptides	
CC	encoded by the polynucleotide is useful for diagnosing human ehrlichiosis	
CC	(Rocky mountain spotted fever) or canine ehrlichiosis. The present	
CC	sequence encodes an E. chaffeensis outer membrane protein.	
XX		
SQ	Sequence 840 BP; 279 A; 128 C; 136 G; 297 T; 0 U; 0 Other;	
	Query Match 20.0%; Score 162.2; DB 9; Length 840;	
	Best Local Similarity 53.8%; Pred. No. 2.1e-28;	
	Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;	
QY	22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTGCTTTTCTGATGCA 81	
Db	34 GTACTTGCATCTCTATTATCATCTTATCTATTGTAATCCTTTTCAGCTATAAATCATAAT 93	
QY	82 AATTTTCTGAAGGAGGAGGAGCACTTTATATAGGTAGTCAGTATATAAGTTGGTATCCC 141	
Db	94 CATACAGGAATAAACAAGTGTGATATATATACAGGCGAGTATACAGCAGGAGTATCC 153	
QY	142 AATTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAGATTTTTCG 201	
Db	154 CATTTTAGCAATTTCTCAGTAAAGAAACAATGTTGATACATACTAGTAGATAT 213	
QY	202 TTAGGCTTTGATAGTCT---GAGATAAATACCTACAGCAATTTTACAGTATCATATGAC 258	
Db	214 AAAAAAAGTCGCTTCTATCGATCCTAACACATTTTCAAACTTTCCAGGTCATATACT 273	
QY	259 CCTACTATGACAGCAGTTTTCAGGGTTTAGTGGTATCATTTGATATTTATGTAATGAC 318	
Db	274 GTTACATTTCAAGATAATGCTGCTAGTTTCAGTGGAGCAATGGATTAITCTTACCCGAA 333	
QY	319 TTT---AGGAGTAAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTAC 375	
Db	334 AGTCTAAGACTTGAAGTTCTTACGAAAAAATTTGATGTCAAAGATCCTAAGAC 393	
QY	376 CCTGAGAATAGCCAAAGCTACAAATTTTTCGCTCGAAATGCTACAAATAGTGAT 435	
Db	394 TACTCAGCAAAAAGATGCTTTTAGGTTTTTTGCTCTAGCACGTAATACGTTCTACTACTGTT 453	
QY	436 AAT-----AGTTTTATAGTACTAGAGATAACGGGTTGTTGACAAGTCTCTTAAT 486	
Db	454 CCTGATGCTCAAAAATATACAGTTATGAAGAAATAAGGCTTATCTGTTGCATCAATCATG 513	
QY	487 GTAAATGTTTGTATGATATTGCTAGTGGTAGTATTCCTTTTAGCACCTTATATATGTTGCT 546	
Db	514 ATCAATGGTTGTATGATCTATCTTTTAATAATTTAGTCGTATACACCTTATATATGCA 573	
QY	547 GGTGTTGTCAGATTATATAAGTTTTTATAGGTATATCATTTGCCCTAAGTTTCTTATCAA 606	
Db	574 GGTATTGGTGAAGATTTCATTGAATTTTGTGACTTTTGCACATTAATAAATTCGCTTATCAA 633	
QY	607 GTTAAGTTTGGTGTCACACTACCTCTAAATGTTTAATACTATGTTGTTGGTGGGGTTAT 666	
Db	634 GGAATACTAGGTATTAGTTATTACTTCTTCTTAAGATTAAATGTTTGGTGGGTGAC 693	
QY	667 TACCATAAGTTGTAGGTAGGATGAGGATGAGAGATAGAAATAGCTTACCATCTACTGCA 726	
Db	694 TATCATAGATTATAGGAATAAATTTAAAAATTT---AATGTTAACCATGTTGTTACA 750	
QY	727 TTATCTGAGTTCTCTAGAACTACTTCAGCTTCTGCTACTTTTAAATPACTGATTATTTGGT 786	
Db	751 CTTGATGAATTTCTTAAAGCAACTTCTGCAGTAGCTACACTTAATGTTGCTTATTTGGT 810	
QY	787 TCGGAGATTGGATTAGATTTC 809	
Db	811 GGTGAAGCTGGAGTAAAGTTTAC 833	
	RESULT 5	
	ADW04249	
ID	ADW04249 standard; DNA; 840 BP.	
XX		
AC	ADW04249;	
DT	24-MAR-2005 (first entry)	
XX		
DE	Ehrlichia chaffeensis outer membrane protein (OMP) DNA, OMP-IV.	
KW	DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;	
KW	infection; vaccine; gene; ds.	
OS	Ehrlichia chaffeensis.	
XX		
FH	Key Location/Qualifiers	
FT	CDS 1..840	
FT	/tag= b	
FT	/product= "Ehrlichia chaffeensis outer membrane protein	
FT	(OMP)"	
FT	sig_peptide 1..78	
FT	/tag= a	
FT	mat_peptide 79..837	
FT	/tag= c	
FT	/product= "Ehrlichia chaffeensis mature OMP protein"	
XX		
PN	US2004265334-A1.	
XX		
PD	30-DEC-2004.	
XX		
PF	29-JUL-2004; 2004US-0090177A.	
XX		
PR	18-SEP-1998; 98US-0100843P.	
PR	19-MAY-1999; 99US-00314701.	
PR	28-JAN-2002; 2002US-00059964.	
XX		
PA	(RIKI/) RIKIHISA Y.	
PA	(OHAS/) OHASHI N.	
XX		
PI	Rikihisa Y, Ohashi N;	
XX		
DR	WPI; 2005-064871/07.	

DR P-PSDB; ADW04250.  
 XX New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia  
 PT canis or Ehrlichia chaffeensis, useful in preparing a composition for  
 PT diagnosing or preventing E. canis or E. chaffeensis infection.  
 XX  
 XX Claim 7; SEQ ID NO 23; 122pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences encoding outer membrane  
 CC proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and  
 CC Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is  
 CC useful in preparing a composition for diagnosing, treating or preventing  
 CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present  
 CC sequence is the Ehrlichia chaffeensis OMP DNA.  
 XX  
 XX Sequence 840 BP; 279 A; 128 C; 136 G; 297 T; 0 U; 0 Other;  
 SQ

Query Match 20.0%; Score 162.2; DB 14; Length 840;  
 Best Local Similarity 53.8%; Pred. No. 2.1e-28;  
 Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTAGCTGAGTACATTTGTTTTTCTTATCTGATGCTGCTTTTCTGATGCA 81  
 DB |||||  
 QY 82 AATTTTCTGAAGGAGGAGGAGCTTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 141  
 DB |||||  
 QY 94 CATACAGGAAATAACACTAGTGGTATATATATACAGGGCAGTATAGACCGAGTATCC 153  
 DB |||||  
 QY 142 AATTTTGTATTTTTCAGCTGAGAACAACTTCTGTATTAACAAAAGATTTTTGCG 201  
 DB |||||  
 QY 154 CATTTTGTAGCAATTTCTCAGTAAAGAACTAATGTTGATACATACACTAGTAGGATAT 213  
 DB |||||  
 QY 202 TTAGTCTTTGATAAGTCT---GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258  
 DB |||||  
 QY 214 AAAAAAGTGGTCTTCTATCGATCTCAACACTTTTCAAGGTCATATACT 273  
 DB |||||  
 QY 259 CCTACTTATGCAAGCAGTTTTCAGGGTTAGTGGTATCATTTGATATATATGTTAATGAC 318  
 DB |||||  
 QY 274 GTTACATTTCAAGATAATGCTGCTAGTTTTCAGTGGAGCAATTTGGATATTTCTTACCCGAA 333  
 DB |||||  
 QY 319 TTT---AGGGTAGAATTTGAGGTTCTATGAGATTTTGAACCTGGAAGCAATGGTAC 375  
 DB |||||  
 QY 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTGATGCTCAAGATCTTAAAGAC 393  
 DB |||||  
 QY 376 CTGAGATAGCCAAAGCTACAAATTTTGTCTGCTCGAAATGCTACAAATAGTGAT 435  
 DB |||||  
 QY 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCACGTAATACGTCTACTACTGTT 453  
 DB |||||  
 QY 436 AAT-----AAGTTTATAGTACTAGAGAAATAACGGCGTTGTGCAAGTCTCTTAAT 486  
 DB |||||  
 QY 454 CTGATGCTCAAAAATATACAGTTATGAAGAATAATGGCTTATCTGTTGCAATCAATCATG 513  
 DB |||||  
 QY 487 GTAATGTTTGTATGATATGCTAGTGGTAGTATCTTTAGCACCTTATATCTGTGCT 546  
 DB |||||  
 QY 514 ATCAATGTTGTATGATCTATCTTTTAATAATTTAGTCGTATCACCTTATATATGTCGA 573  
 DB |||||  
 QY 547 GGTGTTGTGAGATATATAAGTTTTTTTAGTATATCATTTGCGCTTCTTCTTATCAA 606  
 DB |||||  
 QY 574 GGTATGTTGAGATATTCATTTGAAATTTTGTGATCTTTGACATTAACATTTGCTTATCAA 633  
 DB |||||  
 QY 607 GTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAATACTATGTTGTTGGGGGGTTAT 666  
 DB |||||  
 QY 634 GGAAGAACTAGGTATTTAGTTATTTACTTCTTCTAAGATTAATGTTATTTGCTGGGTGAC 693  
 DB |||||  
 QY 667 TACCATAGGTTGTAGGTGATAGGATGAGAGGTAGAAATAGCTTACACTCTACTGCA 726  
 DB |||||  
 QY 694 TATCATAGAGTTATAGGGAATAAATTTAAAAATTT---AAATGTTTAAACCATGTTGTACA 750  
 DB |||||  
 QY 727 TTATCTGAGGTTCTCTAGAACTACTTTCAGCTTCTGCTACTTCTTAAATACTGATTATTTTGGT 786  
 DB |||||  
 QY 751 CTTGATGAATTTCTTAAAGCAACTTCTGAGTAGCTACATTAATGTTGCTTATTTTGGT 810  
 DB |||||

QY 787 TCGGAGATTGGATTAGATTTC 809  
 DB 811 GGTGAAGCTGGAGTAAAGTTTAC 833

RESULT 6  
 AAX34755  
 ID AAX34755 standard; DNA; 852 BP.  
 XX AAX34755;  
 AC  
 XX 27-AUG-2003 (revised)  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE DNA encoding OMP-1W protein.  
 XX  
 XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;  
 KW detection; dog; ss.  
 XX  
 XX Ehrlichia chaffeensis.  
 OS  
 XX WO9913720-A1.  
 XX  
 XX 25-MAR-1999.  
 XX  
 PF 18-SEP-1998; 98WO-US019600.  
 XX  
 XX 19-SEP-1997; 97US-0059353P.  
 XX  
 XX (OHIS ) UNIV OHIO STATE.  
 XX  
 XX Rikihisa Y, Ohashi N;  
 PI  
 XX WPI; 1999-254290/21.  
 DR  
 DR P-PSDB; AAY06955.  
 XX  
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia  
 PT canis.  
 XX  
 XX Disclosure; Fig 15A; 55pp; English.  
 XX  
 CC The invention provides isolated outer membrane proteins (OMP) from  
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part  
 CC of the OMP family and consist of proteins OMP-1, -11B to Z) shown in  
 CC AAY06943-958. The E. canis proteins form part of the P30 family and  
 CC consist of proteins shown in AAY06959-970. The proteins and genes are  
 CC used to detect E. chaffeensis in patients and E. canis in dogs. (Updated  
 CC on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 U; 0 Other;  
 SQ

Query Match 18.5%; Score 150.8; DB 2; Length 852;  
 Best Local Similarity 54.6%; Pred. No. 1e-25;  
 Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

QY 96 GAGAGAGGACTTTATATAGGTAGTCAGTATAAAGTTGGTATTCCTCAATTTTAGTAATTT 155  
 DB |||||  
 QY 111 GTGCACCTGGCGCTATATGTCAGTGCACAAATATAAACCTACTGTTTCTCCTTTAGTAATTT 170  
 DB |||||  
 QY 156 TTCAGCTGAGAGAACAAATTCCTGATTTACAAAAGATTTTTCGCTTAGCTCTTGATAA 215  
 DB |||||  
 QY 171 TTCACCTTAAAGAAACTTATATCTGACACTAAAGAGTTATTAGGACTACGAAAGATATTA 230  
 DB |||||  
 QY 216 GTCT-----GAGATAAATACTCACAGCAATTTTACACGATCATATGACCTTACTTATGC 269  
 DB |||||  
 QY 231 GTCTATTACAGATATAACAATAAATAAATTCACATTTCTTATAACACAAATTTCA 290  
 DB |||||  
 QY 270 AAGCAGTTTTCAGGGTTTATAGTGTATCATTTGATATATTTGTAATGACTTT---ACGGT 326  
 DB |||||  
 QY 291 AGATAATGCTGTTAGCTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGT 350  
 DB |||||  
 QY 327 AGAATTTGAGGTTCTTATGAGATTTTGAACCTGGAAGCAATGGTACCTGAGATAG 386  
 DB |||||

Db 351 TGAGGTAGATGCTTATGAGAAATTGACGTTAAAAATCCTGGTAATTACGTAGTAAG 410  
 Qy 387 CCAAGCTCAAAATTTTTCCTTCTCGAATGCTACAAATAGTATGAT----- 435  
 Db 411 TGAAGCCTTCAGGTATATGCTTTAGCAAGAGAAATTGATAATCTTCAAAAATATCCTGA 470  
 Qy 436 ----AATAAGTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAGTCTCTTAATGAAA 491  
 Db 471 AACAAATAGTATGTTGTTATAAGAACAAATGCTTATCTGTGCAATCCATTAATCAA 530  
 Qy 492 TGTGTTGTTATGATATGCTAGTGGTATGATTCCTTTAGCACCTTATATGCTGGTGT 551  
 Db 531 TGGCTGTATGATTTTCTTTAAACAATTAAAGATATCACCTTACATATGCTAGGGTT 590  
 Qy 552 TGGTCAGATATATAAAGTTTTTATGATATCAATGCTTATGCTTATCAAGTTAA 611  
 Db 591 TGGTGGGCAATATATAGAAATTTTATGCTGCTGAAGTTTAAATTTGCTTATCAAGTAA 650  
 Qy 612 GTTGTGTCAACTACCTCTAAATGTTAATATCTATGTTGTTTGGGGGTTTATACCA 671  
 Db 651 GGTAGTATCAGTTATCCATTTCTTAATATGATATATTTGCTGACGGATATACCA 710  
 Qy 672 TAAGTGTAGGTAGTAGGATGAGAGATGAGAAATAGCTTACCATCTTCAATGATC 731  
 Db 711 TAAGGTATAGGAAATAAATTTAAACAATTTAAATGTTCAACACGTTGTAGTCTTAACAG 770  
 Qy 732 TGAGGTTCTAGAACTTCTAGCTTCTGCTACTTTAAATATGATATTTGGTGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTGCAAGTACTCTTAAATGTTGAGTATTTCCGTAGTGA 827  
 Qy 792 GATTGGATTTAGATTT 807  
 Db 828 ATTGGGTTAAATTT 843

RESULT 7  
 ABS63288  
 ID ABS63288 standard; DNA; 852 BP.  
 XX  
 AC ABS63288;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE DNA encoding Ehrlichia chaffeensis outer membrane protein (OMP) #13.  
 XX  
 KW Outer membrane protein; OMP; gene; ds; P30F; ehrlichiosis; infection.  
 XX  
 OS Ehrlichia chaffeensis.  
 XX  
 PN US2002120115-A1.  
 XX  
 PD 29-AUG-2002.  
 XX  
 PF 28-JAN-2002; 2002US-00059964.  
 XX  
 PR 19-MAY-1999; 99US-00314701.  
 XX  
 PA (RIKI/) RIKIHISA Y.  
 PA (OHAS/) OHASHI N.  
 XX  
 PI Rikihisa Y, Ohashi N;  
 XX  
 DR WPI; 2002-618954/66.  
 DR P-PSDB; ABG77947.  
 XX  
 PT Isolated polynucleotide encoding an outer membrane protein of E.canis or  
 E.chaffeensis used in the diagnosis of infection.  
 XX  
 PS Claim 3; Fig 15A; 49pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide encoding an outer  
 CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
 CC in the diagnosis of infection. An infection such as human ehrlichiosis or

CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
 CC patient, providing a polypeptide or mixture of polypeptides, contacting  
 CC the sample with the polypeptide and assaying for the formation of a  
 CC complex between antibodies in the serum sample and the polypeptide, where  
 CC formation of a complex is indicative of infection with E. chaffeensis.  
 CC This sequence represents DNA encoding an Ehrlichia outer membrane protein  
 CC of the invention  
 XX  
 SQ Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 U; 0 Other;  
 Query Match 18.5%; Score 150.8; DB 6; Length 852;  
 Best Local Similarity 54.6%; Pred. No. 1e-25;  
 Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;  
 Qy 96 GAGGAGAGGACTTTATATAGTAGTACAGTATATAAGTTGGTATTCCTCAATTTAGTAATTT 155  
 Db 111 GTGCACTGGGCTATATGTCAGTGGACAATATAAACCTACTGTTCTCTCACTTTAGTAATTT 170  
 Qy 156 TTCAGCTGAAGAAACAAATTCCTGGTATTACAAAAAAGATTTTGGCTTAGGCTCTTGATAA 215  
 Db 171 TTCACCTTAAGAACTTATCTGACACTTAAGAGTTATTAGGACTAGCAAAAGATATTAA 230  
 Qy 216 GTCT-----GAGATAAATACTCAGCAATTTTACAGATCATATGACCCCTACTTATGC 269  
 Db 231 GTCTATTACAGATATAACAACAATAAAAAATTCACATTCCTTATAACAAAAATTTCA 290  
 Qy 270 AAGCAGTTTTGCNGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTT--AGGGT 326  
 Db 291 AGATAATGCTGTTAGCTTCAGTGCAGCTGTTGGATATATTTCCCAAGACACGTCGAAGGT 350  
 Qy 327 AGAATTTGAAGGTTCTTATGAGAAATTTGAACCTGAAAGACAATGGTACCCTCGAGAATAG 386  
 Db 351 TGAGGTAGATGGTCTTATGAGAAATTTGACGTTAAATAATCCTGTATATTAGTAGTAAG 410  
 Qy 387 CAAAAGCTACAAATTTTGTCTCGAAATGCTACAAATAGTGAT----- 435  
 Db 411 TGAAGCCTTCAGGTATATGCTTTAGCAAGAGGAATGATAATCTTCAAAAATATCCTGA 470  
 Qy 436 ----AATAAGTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAGTCTCTTAATGTA 491  
 Db 471 AACAAATAGTATGTTGTTATAAAGAACAAATGGCTTATCTGTGCAATCCATTAATCAA 530  
 Qy 492 TGTGTTGTTATGATATGCTAGTGGTATGATTCCTTTAGCACCTTATATGCTGGTGT 551  
 Db 531 TGGCTGTATGATTTTCTTTAAACAATTTAAAGATACACCTTACATATGCTAGGGTT 590  
 Qy 552 TGGTCAGATATATAAAGTTTTTATGATATCATTTGCTAAGTTTCTTATCAAGTTAA 611  
 Db 591 TGGTGGGACATTAAGAATTTTGTAGTGTGTAAGTTTAAATTTGCTTATCAAGGTA 650  
 Qy 612 GTTGGTGTCAACTACCTCTAAATGTTAATACTATGTTGTTGGGGGTTATACCA 671  
 Db 651 GGTAGGTATCAGTTATCCATTTATCTTAATGATATATTTGCTGACGGATATACCA 710  
 Qy 672 TAAGGTGTAGGTAGTATGAGAGATGAGAAATAGCTTACCATCTTCTGCTGCTATATC 731  
 Db 711 TAAGGTATAGGAAATAAATTTAAACAATTTAAATGCTTCAACGTTGTTAGTCTTAACAG 770  
 Qy 732 TGACGTTCTAGAACTATCTGAGCTTCTGCTACTTTTAAATATGATATTTGGTGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTTGCAGTAGTACTCTTAAATGTTGAGTATTTCCGTAGTGA 827  
 Qy 792 GATTGGATTTAGATTT 807  
 Db 828 ATTGGGTTAAATTT 843  
 RESULT 8  
 ADA09758  
 ID ADA09758 standard; DNA; 852 BP.  
 XX  
 AC ADA09758;  
 XX





DR WPI; 2005-064871/07.  
 DR P-PSDB; ADW04252.  
 XX New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia  
 PT canis or Ehrlichia chaffeensis, useful in preparing a composition for  
 PT diagnosing or preventing E. canis or E. chaffeensis infection.  
 XX  
 PS Claim 7; SEQ ID NO 25; 122pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences encoding outer membrane  
 CC proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and  
 CC Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is  
 CC useful in preparing a composition for diagnosing, treating or preventing  
 CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present  
 CC sequence is the Ehrlichia chaffeensis OMP DNA.  
 XX  
 SQ Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 U; 0 Other;  
 Query Match 18.5%; Score 150.8; DB 14; Length 852;  
 Best Local Similarity 54.6%; Pred. No. 1e-25;  
 Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;  
 QY 96 GAGGAGGAGCTTTATATAGGTAGTCAGTATAAAGTTGGTATCCCAATTTTAGTAAATTT 155  
 Db 111 GTGCACCTGGGCTATATGTCACTGAGCAATATAAACCCTACTGTTCTCACTTTAGTAAATTT 170  
 QY 156 TTCAGCTGAGAACAATCTCTGGTATTACAAAAAGATTTTTCGTTAGTCTTGATAA 215  
 Db 171 TTCACCTAAAGAACTTATATACACACTAAGAGTTATTAGGACTAGCAAAAAGATATTA 230  
 QY 216 GTCT-----GAGATAAATCTCACAGCAATTTTACAGCATCATATGACCTACTTATGC 269  
 Db 231 GTCTATTACAGATATACACAATAAATAATTCACATTCCTTATACACAATAATTTCA 290  
 QY 270 AAGCAGTTTTCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTT---AGGGT 326  
 Db 291 AGATAATGCTGTTAGCTTCAGTCAGCTGTTGGATATATTTCCTCAAGACAGTCCAGGGT 350  
 QY 327 AGAATTTGAAGTCTTATGAGAAATTTTGACCTGAAAGCAATGTTACCTGAGAAATAG 386  
 Db 351 TGAGGTAGAAATGCTCTTATGAAGAATTTGACGTTAAATAATCTCGGTAATTTACGTAGTA 410  
 QY 387 CCAAGCTCAAAATTTTTCCTTGTCTCGAAATGCTACAAATAGTGAT----- 435  
 Db 411 TGAAGCTTCAGGTATATGCTTTTAGCAAGAGAAATGATAATCTTCAAAAATATCCTGA 470  
 QY 436 ----AATAAGTTATAGTACTAGAGAAATAACGCGTTGTTGACAACTCTCTTAATGTAAA 491  
 Db 471 AACAAATAAGTATGTTGTTATAAAGAACAAATGCTTATCTGCGCATCCATTATAATCAA 530  
 QY 492 TGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCTGGTGT 551  
 Db 531 TGCTGTATGATATTTTCTTTAAACAATTTTAAAGATATCACCTTACATATGCGTAGGGTT 590  
 QY 552 TGCTGCAGATATATAAAGTTTTCAGTATATCATTTGCCCTAAGCTTTTCTTATCAAGTAA 611  
 Db 591 TGTTGGGACATATAGAAATTTTTCAGTGTGCTGTAAGTTTAAATTTGCTTATCAAGTAA 650  
 QY 612 GTTTTGTGCTCAACTACCTCTAAATGTTAATATCTATGTTGTTGGGGGTTTATTAACA 671  
 Db 651 GGTAGGTATCAGTTATCCATTATCTCTATATATGATTATATTGCTGACGATATTAACA 710  
 QY 672 TAAGTTGTAGTGATAGGATGAGAGATAGAAATAGCTTACCATCCTACTGCATTATC 731  
 Db 711 TAAGGTGATAGAAATTAATTAACAAATTTAAATGTTCAACAGCTTGTAGTCTTAACAG 770  
 QY 732 TGACGTTCTCAGAACTACTCAGCTTCTGCTACTCTTAAATACATGATTTTGGTTGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTTGAGTAGTACTCTTAAATGTTGAGTATTTTCGTTAGTGA 827  
 QY 792 GATTGGAATTTAGATTT 807  
 Db 828 ATTTGGGTTAAAAATTT 843

RESULT 10  
 ABK68876  
 ID ABK68876 standard; DNA; 840 BP.  
 XX  
 AC ABK68876;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE DNA encoding Ehrlichia canis p28-2.  
 XX  
 KW Ehrlichia canis infection; vaccine; serodiagnostic; gene; p28; ss;  
 KW antibacterial.  
 XX  
 OS Ehrlichia canis.  
 XX  
 PN WO200222782-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 12-SEP-2001; 2001WO-US028759.  
 XX  
 PR 12-SEP-2000; 2000US-00660587.  
 XX  
 PA (RERE-) RES DEV FOUND.  
 XX  
 PI Walker DH, Yu X, McBride JW;  
 DR WPI; 2002-351882/38.  
 DR P-PSDB; AAU96116.  
 XX  
 PT New recombinant homologous 28 kilodalton immunodominant protein from  
 PT Ehrlichia canis, useful for treating Ehrlichia canis infections.  
 XX  
 PS Claim 5; Fig 14; 106pp; English.  
 XX  
 CC The invention relates to a recombinant homologous 28 kDa immunodominant  
 CC protein, p28 (I) of Ehrlichia canis. (I), a 28-kDa antigen preferably  
 CC dispersed in a pharmaceutically acceptable carrier, is useful for  
 CC inhibiting E. canis infection in a subject. (I) is useful in the  
 CC development of vaccines and serodiagnostics that are particularly  
 CC effective for disease prevention and serodiagnosis. ABK68852-ABK68878  
 CC represent the 28-kDa antigen coding sequences and PCR primers of the  
 CC invention  
 XX  
 SQ Sequence 840 BP; 301 A; 141 C; 144 G; 254 T; 0 U; 0 Other;  
 Query Match 18.1%; Score 147.2; DB 6; Length 840;  
 Best Local Similarity 53.2%; Pred. No. 7.4e-25;  
 Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;  
 QY 1 ATGAATTAACAAAGATTTGTTGAGTGTACGCTGAGTACATTTGTTTTTCTTATCT 60  
 Db 1 ATGAATTAACAAAGATTTGTTGAGTGTACGCTGAGTGTACATTTGTTTTTCTTATCT 60  
 QY 61 GATGCTGCTTTTCTGA-----TGCATATTTTCTGAGGAGGAGGAGCTTTAT 111  
 Db 61 TATCAGTCTTTTGCAGATCCTGTAGGTTCAAGAACTAATGATTAACAAAGAGGCTTCTAC 120  
 QY 112 ATAGGTAGTCAGTATAAGCTTGGTATTCCCAATTTTGTAGTAAATTTTTCAGCTCAAGAAACA 171  
 Db 121 ATTAGTCAAGAGTACAAATCCCAAGTATATCACATTTAGAAAATTTCTCTGCTGAGAACT 180  
 QY 172 ATCTCTGGTATTAACAAAAAGATTTTGGTTAGTCTTTG-----ATAAGTCTGAG 222  
 Db 181 CCTATTAATGGAACAAATTTCTCTACTAAAAAAGTTTTCGGACTAAAGAAAGATGGTAT 240  
 QY 223 ATAATATCTCACAGCAATTTTACAGATCATATGACCTACTTAT-----GCAAGCACT 276  
 Db 241 ATAAACAAAAAAGACGATTTTCAAGAGTAGTCCAGGCAATTTGATTTTCAAAATCACTTA 300  
 QY 277 TTTGCGGGTTTGTAGTGTATCATTTGATATTATTTAATGACTTTTAGGTAGAAATTTGAA 336

```
Db 301 ATATCAGGATTTTCAGGAAGTATGGTTACTATGACGACCAAGATAGAACTTGAA 360
Qy 337 GGTTCCTTATGAGAAATTTGAACTTGAAGACAATGGTACC-----CTGAGAATAGCAA 390
Db 361 GCTGCATATCAACAATTTAATCCAAAAACACCGATAAACAATGATATGATAATGGTGAA 420
Qy 391 AGCTACAAAATTTTGTGCTTCGAAATGCTCAAAATAGTGAATAAAGTTTATAGTA 450
Db 421 TACTATATAACATTTTGCATTTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480
Qy 451 CTACAGAAATACGCGGTGTTGCAAGCTCTTAATGTAATGTTTCTGTTATGATTTGCT 510
Db 481 CTTAAATAGCGGCAATCTTTATGTCATGATGTTTATGATGTTTATGATGTTTATGAT 540
Qy 511 AGTGGTAGTATTCCTTTAGCACCTTATATGTCGCTGCTGTTGTTGTCGAGATTTATATAAG 570
Db 541 GCTGAAGGATATCTTTTCGTACCATATGTCATGTCAGGTATAGGAGCAGATCTTATCACT 600
Qy 571 TTTTATGATATATCAATGCGCTTAAAGTTTCTTATCAAGTTAAGTTTGGTGTCACTACCT 630
Db 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTACCAGGAAAAATAGGTATTTAGTTACCT 660
Qy 631 CTAAATGTTAATACTATGTTGTTGGTGGGTTTATACCATAGGTTTGTAGGTATAGG 690
Db 661 ATCACACAGAAAGTCTCTGCAATTTATGTTGGTGGATACCTACCATGCGGTTATGGTAATAA 720
Qy 691 CATGAGAGATAGAAATAGCTTACCATCTCTGATTTATGTCAGTTTCTGAGAACTACT 750
Db 721 TTTGAGAAGATACCTGTA---ATAACTCTCTGTAGTATTTAAATGATGCTCTCTCAACCCACA 777
Qy 751 TCAGCTTCTGCTACTTTTAAATACGTATTTTGGTTGGGAGATTTGAAATTTAGATT 806
Db 778 TCTGCTTCAGTAACCTCTTGACGTTGGATCTTTGGCGGAGAAATTTGGAATGAGGTT 833
```

## RESULT 11

```
ABS63299
ID ABS63299 standard; DNA; 843 BP.
AC ABS63299;
XX
DT 15-NOV-2002 (first entry)
DE DNA encoding Ehrlichia canis outer membrane protein (P30F) #9.
KW Outer membrane protein; OMP; gene; ds; P30F; ehrlichiosis; infection.
XX
OS Ehrlichia canis.
XX
PN US2002120115-A1.
XX
PD 29-AUG-2002.
XX
PF 28-JAN-2002; 2002US-00059964.
XX
PR 19-MAY-1999; 99US-00314701.
XX
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2002-618954/66.
DR P-PSDB; ABG77958.
XX
PT Isolated polynucleotide encoding an outer membrane protein of E. canis or
PT E. chaffeensis used in the diagnosis of infection.
XX
PS Claim 5; Fig 30A; 49pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding an outer
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
```

```
CC in the diagnosis of infection. An infection such as human ehrlichiosis or
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC patient, providing a polypeptide or mixture of polypeptides, contacting
CC the sample with the polypeptide and assaying for the formation of a
CC complex between antibodies in the serum sample and the polypeptide, where
CC formation of a complex is indicative of infection with E. chaffeensis.
CC This sequence represents DNA encoding an Ehrlichia outer membrane protein
CC of the invention
XX
SQ Sequence 843 BP; 303 A; 141 C; 144 G; 255 T; 0 U; 0 Other;
```

```
Query Match 18.1%; Score 147.2; DB 6; Length 843;
Best Local Similarity 53.2%; Pred. No. 7.4e-25;
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;

Qy 1 ATGAATATACAAAGATTTGTGTGAGTGTACCTGAGTACATTTGTTTTTCTTATCT 60
Db 1 ATGAATATATAAGAAAATTTCTAGTAAGAGCGCTTAATCTCATTAATGTCAATCTTACCA 60
Qy 61 GATGGTCTTTTCTGA-----TGCAAAATTTTCTGAAGGAGGAGGACTTTAT 111
Db 61 TATCAGTCTTTTGCAGATCTCTGAGTTCAAGAACTAAATGATAACAAAGAGGCTTCTAC 120
Qy 112 ATAGGTAGTCAGTATAAAGTTGGTATTTCCCAATTTTAGTAATTTTTCAGCTGGAAGAAACA 171
Db 121 ATTAGTGCAAGTACAACTCCAGTATATCACCTTTAGAAAATTTCTCTGCTGGAAGAACT 180
Qy 172 ATTCTCTGGTATTAACAAAAAGATTTTTCGCTTAGTCTTG-----ATAAGTCTGAG 222
Db 181 CCTATTAAATGGAACAAATCTCTCACTAAAAAGTTTTCGGACTAAAGAAAGATGGTGAT 240
Qy 223 ATAAATCTCACAGCAATTTTACAGTACATATGACCCCTACTTAT-----GCAAGCAGT 276
Db 241 ATAACAAAAAAGACGATTTTACAAGAGTAGTCCAGGCAATGATTTTCAAAAATAACTTA 300
Qy 277 TTTGCGAGGTTTGTAGTGTATCATTTGGATATTTATGTTAATGACTTTAGGGTAGAATTTGAA 336
Db 301 ATATCAGGATTTTCAGGAAGTATTTGGTACTCTATGACGCGCAAGAAATAGAACTTGAA 360
Qy 337 GGTTCCTTATGAGAAATTTTGAACCTGAAAGACAATGGTACC-----CTGAGAATAGCCAA 390
Db 361 GCTGCATATCAACAATTTTAAATCCAAAAAACACCGATAACAATGATGATACTGATGATGAA 420
Qy 391 AGCTACAAAATTTTGTCTTCTCGAAATGCTACAAATAGTATGATAAAGTTTATAGTA 450
Db 421 TACTATAAACAATTTTGCATTTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480
Qy 451 CTAGAGAAATACGCGGTGTTGACAAGTCTCTTAATGTAATGTTTCTGTTATGATATTGCT 510
Db 481 CTTAAAAATGACGCGCATAACTTTTATGCTATTGATGTTTAATCTCTGATGACATTACA 540
Qy 511 AGTGGTAGTATTCCTTTAGCACCTTATATGTCGCTGCTGTTGGTGCAGATTTATATAAG 570
Db 541 GCTGAAGGAGTATCTTTTCGTACCATATGTCATGTCAGGTATAGGAGCAGATCTTATCACT 600
Qy 571 TTTTATGATATATCATTTGCTTAAAGTTTCTTATCAAGTTAAGTTTGGTGTCACTACCT 630
Db 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTACCAGGAAAAATAGGTATTTAGTTACCT 660
Qy 631 CTAAATGTTAATACTATGTTTGGTGGGTTTATACCATAGGTTTGTAGGTATAGG 690
Db 661 ATCACACAGAAAGTCTCTGCAATTTATGTTGGTGGATACCTACCATGCGGTTATGGTAATAA 720
Qy 691 CATGAGAGATAGAAATAGCTTACCATCTCTGATTTATGTCAGTTTCTGAGAACTACT 750
Db 721 TTTGAGAAGATACCTGTA---ATAACTCTCTGTAGTATTTAAATGATGCTCTCTCAACCCACA 777
Qy 751 TCAGCTTCTGCTACTTTTAAATACGTATTTTGGTTGGGAGATTTGAAATTTAGATT 806
Db 778 TCTGCTTCAGTAACCTCTTGACGTTGGATCTTTGGCGGAGAAATTTGGAATGAGGTT 833
```

RESULT 12

[illegible]

```
PR 19-MAY-1999; 99US-00314701.
PR 28-JAN-2002; 2002US-00059964.
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
XX Rikihisa Y, Ohashi N;
XX
XX WPI; 2005-064871/07.
XX P-PSDB; ADW04274.
XX
XX New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia
XX canis or Ehrlichia chaffeensis, useful in preparing a composition for
XX diagnosing or preventing E. canis or E. chaffeensis infection.
XX
XX Claim 2; SEQ ID NO 47; 122pp; English.
XX
XX The invention relates to nucleic acid sequences encoding outer membrane
XX proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and
XX Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is
XX useful in preparing a composition for diagnosing, treating or preventing
XX an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present
XX sequence is the Ehrlichia canis P30F DNA.
XX
XX Sequence 843 BP; 303 A; 141 C; 144 G; 255 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 147.2; DB 14; Length 843;
XX Best Local Similarity 53.2%; Pred. No. 7.4e-25;
XX Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;
XX
QY 1 ATGAATTACAAAGATTGTTGGTAGGTGTACCGTGGAGTACATTTGTTTCTTATCT 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 ATGAATTATAGAAAATTTCTAGTAAGAAGCGGTTAATCTCATTTAATGTCAATCTTACCA 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GATGGTCTTTTCTGA-----TGCRAATTTTCTGAAGGAGGAGGAGCTTTAT 111
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 TATCAGTCTTTTGAGATCCGTGATGTTCAAGAATAATGATATAACAAGAGGCTTCTAC 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 112 ATAGGTAGTCAGTATAAAGTTGGTATTCCTCAATTTTATAGTAAATTTTTCAGCTGAAGAAACA 171
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ATTAGTGCAGATACATCCAAAGTATATACACATTTAGAAAATTTCTCTGCTGAAGAAACT 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 172 ATTCCTGGTATACAAAAGATTTTTCGGTGTAGTCTTG-----ATAAGTCTGAG 222
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CCTATTAAATGCAAAATTTCTCTCACTAAAAAGTTTTCGGACTAAAGAAAGATGGTAT 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 223 ATAAATCTCAGCAATTTTACAGCATCATAGCCCTACTAT-----GCAAGCAGT 276
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 ATAAACAAAAGACGATTTTACAGAGTAGCTCCAGGCATGATTTTCAAAATAACTTA 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 TTTCGAGGTTTGTAGGTATCATTTGGATATATTATGATGACTTTTAGGGTAGAATTTGAA 336
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 ATATCAGGATTTTCAGGAAGTATTTGTTACTCTATGGAGGACCAAGATAGACTTGA 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 GGTCTTATAGAAATTTTGAACCTGGAAGACAAATGGTACC-----CTGAGAAATAGCCAA 390
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GCTGCATATCAACAAATTTAATCCAAAACCAACCGATACATGATGATAAATGGTGAA 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 391 AGCTACAAATTTTGTCTTCTCGAATGCTCAATAGTCAATAGTATATAGTTATAGTA 450
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 TACTATAAACATTTTGCATATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 451 CTAGAGAAATAACGGGCTTGTGACAAAGTCTCTTAATGTAATGTTTGTATGATATGCT 510
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 CTTAAATATGCGGCATAACTTTTATGTCATTGATGTTTATACTTGTATGACATTACA 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 511 AGTGATGATATCTCTTTAGACCTTATATGTGTGCTGTGTGTGTGCGAGATATATAAG 570
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 GCTGAAGGAGTATCTCTTGTACCATATGCATGTGCAGGTATAGGACGAGATCTTATCACT 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 571 TTTTATAGGTATATCATTCCTTAAGTTTCTTATCAAGTTAAGTTTGGTGCACTACCT 630
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 ATTTTAAAGACCTCAATCTAAATTTTGTCTTACCAGGAAATAAGGTATTAGTTTACCCT 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 631 CTAATGTTAATACTATGTTGTTGGTGGGGTTATTACCAATAAGCTTGTAGGTATAGG 690
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 ATCACACCAAGAGTCTCTGCAATTTATTGGTGGATACCATGGCGTTATTGGTAATAAA 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 691 CATGAGAGAGTAGAAATAGCTTACCATCTTACTGCAATTTATCTGACGTTCTAGAACCTACT 750
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 TTTGAGAGATACCTGTA---ATACTCTCTGTAGTATTAAATGATGCTCTCTCAACACCA 777
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 751 TCAGCTTCTGCTACTTTAAATACTGAATTTTGGTTGGGAGATTGGATTGATT 806
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 TCTGCTTTCAGTAATCTTTCAGCGTTGGATACTTTGGCGGAGAAATTGGAATGAGTT 833
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 14
ABS63306
ID ABS63306 standard; DNA; 726 BP.
XX
XX ABS63306;
XX AC
XX 15-NOV-2002 (first entry)
XX DT
XX DNA encoding Ehrlichia canis outer membrane protein (P30F) #14.
XX DE
XX Outer membrane protein; OMP; gene; ds; P30F; ehrlichiosis; infection.
XX KW
XX Ehrlichia canis.
XX OS
XX US2002120115-A1.
XX PN
XX 29-AUG-2002.
XX PD
XX 28-JAN-2002; 2002US-00059964.
XX PF
XX 19-MAY-1999; 99US-00314701.
XX PR
XX (RIKI/) RIKIHISA Y.
XX PA (OHAS/) OHASHI N.
XX
XX Rikihisa Y, Ohashi N;
XX PI
XX WPI; 2002-618954/66.
XX DR P-PSDB; ABG77965.
XX
XX Isolated polynucleotide encoding an outer membrane protein of E. canis or
XX E. chaffeensis used in the diagnosis of infection.
XX
XX Disclosure; Fig 32A; 49pp; English.
XX
XX The invention relates to an isolated polynucleotide encoding an outer
XX membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
XX in the diagnosis of infection. An infection such as human ehrlichiosis or
XX canine ehrlichiosis can be diagnosed by providing a serum sample from the
XX patient, providing a polypeptide or mixture of polypeptides, contacting
XX the sample with the polypeptide and assaying for the formation of a
XX complex between antibodies in the serum sample and the polypeptide, where
XX formation of a complex is indicative of infection with E. chaffeensis.
XX This sequence represents DNA encoding an Ehrlichia outer membrane protein
XX of the invention
XX
XX Sequence 726 BP; 255 A; 99 C; 112 G; 260 T; 0 U; 0 Other;
XX
XX Query Match 18.0%; Score 146.6; DB 6; Length 726;
XX Best Local Similarity 54.0%; Pred. No. 9.8e-25;
XX Matches 358; Conservative 0; Mismatches 284; Indels 21; Gaps 2;
XX
QY 166 GAACAATTCCTGGTATTACAAAAGATTTTTCGGTTAGTCTTGTAGTCTGAGATA 225
DB 55 GAGGCATTTGTTGGCTTAAACAAGATATTAGTTCTATTTTACGTAATAAGAGACCACA 114
QY 226 AATACTCACAGCAATTTTACACGATCATATGACCCCTACTTATGCAAGCAGTTTTCGAGG 285
DB 115 CAATATAATAACAATTTTAAAGTTCCTCTATCTGCAAAATTTCAAGAGGACTTTGCGAGT 174
```



Db 715 TTT 717

Search completed: December 28, 2005, 10:56:44  
Job time : 626 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 10:46:12 ; Search time 4129 Seconds  
(without alignments)  
9212.375 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaagaattgt.....ttggatttagattgcgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_est3.\*
- 4: gb\_est4.\*
- 5: gb\_est5.\*
- 6: gb\_est6.\*
- 7: gb\_est7.\*
- 8: gb\_est8.\*
- 9: gb\_gsa1.\*
- 10: gb\_gsa2.\*
- 11: gb\_gsa3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62.8	7.7	1101	10	CNS0039G	AL063921 Drosophil
2	60.8	7.5	1101	10	CNS0100X	AL098379 Drosophil
3	56.8	7.0	829	10	EX173672	EX173672 Danio rer
4	56.4	6.9	1101	10	CNS00EVL	AL069706 Drosophil
5	55.8	6.9	1101	10	CNS00ESI	AL069797 Drosophil
6	54.2	6.7	1101	10	CNS00LOO	AL068607 Drosophil
7	54	6.6	530	7	CK573092	CK573092 1099 Plas
8	53	6.5	1101	10	CNS00LT2	AL078714 Drosophil
9	52.6	6.5	976	10	CNS01ODV	AL098845 Drosophil
10	51.2	6.3	630	10	EX152436	EX152436 Danio rer
11	51	6.3	844	10	EX139987	EX139987 Danio rer
12	51	6.3	889	10	CK973509	CK973509 AIAA-aaa5
13	50.4	6.2	1094	10	CNS012PZ	AL101513 Drosophil
14	50.2	6.2	471	10	CG807194	CG807194 1118078H0
15	50	6.2	670	9	BZ388480	BZ388480 EINDL21TF
16	49.8	6.1	429	10	CG803197	CG803197 1118039H0
17	49.6	6.1	753	5	BW073695	BW073695 BW073695
18	49.6	6.1	791	5	BW143210	BW143210 BW143210
19	49.6	6.1	930	9	AZ549806	AZ549806 ENTFL24TR
20	49.4	6.1	1068	10	CNS00J4X	AL075933 Drosophil
21	49.2	6.1	641	9	AQ946120	AQ946120 Sheared D
22	49.2	6.1	1101	10	CNS00FMC	AL070972 Drosophil

23	49.2	6.1	1131	11	CNS034FO	AL227373 Tetraodon
24	49	6.0	878	10	CNS0187R	AL108993 Drosophil
25	49	6.0	1101	10	CNS0176D	AL107647 Drosophil
26	48.6	6.0	891	10	CW952279	CW952279 TcB38.1LB
27	48.4	6.0	1101	10	CNS00807	AL069440 Drosophil
28	48.2	5.9	1183	1	AJ926017	AJ926017 AJ926017
29	48	5.9	811	8	CK092809	CK092809 EHAFJ16TR
30	47.8	5.9	735	10	CW960936	CW960936 AIAA-aacl
31	47.8	5.9	767	10	EX193880	EX193880 Danio rer
32	47.6	5.9	332	5	BW139522	BW139522 BW139522
33	47.6	5.9	739	5	BW257592	BW257592 BW257592
34	47.6	5.9	783	5	BW131201	BW131201 BW131201
35	47.6	5.9	785	5	BW127705	BW127705 BW127705
36	47.4	5.8	320	1	AV984612	AV984612 AV984612
37	47.4	5.8	352	5	BW269652	BW269652 BW269652
38	47.4	5.8	354	5	BW269352	BW269352 BW269352
39	47.4	5.8	478	5	BW262060	BW262060 BW262060
40	47.4	5.8	496	3	BP018342	BP018342 BP018342
41	47.4	5.8	575	5	BW430026	BW430026 BW430026
42	47.4	5.8	713	5	BW480796	BW480796 BW480796
43	47.4	5.8	987	10	CNS014FQ	AL104456 Drosophil
44	47.2	5.8	580	11	CR479976	CR479976 mCh2-190L
45	47.2	5.8	685	10	CG957658	CG957658 MBEEB62TR

ALIGNMENTS

RESULT 1  
CNS0039G/c

LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL063921  
VERSION  
AL063921.1 GI:4941778  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 1101)  
Genoscope.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
JOURNAL  
- Web : www.genoscope.cns.fr)  
COMMENT  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoer in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's Pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source

1. .1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR08K10"  
/clone\_lib="RPCI-98"  
/notes="end : TET3"

ORIGIN

```
Query Match 7.7%; Score 62.8; DB 10; Length 1101;
Best Local Similarity 15.8%; Pred. No. 0.00051;
Matches 102; Conservative 272; Mismatches 270; Indels 0; Gaps 0;

QY 11 AAGATTGTTAGTGGTTCACCTGAGTACATTTCTTTTCTTATCTGATGGTCTT 70
DB 1061 DADRWAADDGAGTWTATWMMWMAWTDWTDKMMWATAAKTDTAWTWRATAWADW 1002

QY 71 TTTCTGATCAAAATTTTCTGAAGGAGGAGGACTTTATATAGTAGTCAATTAAG 130
DB 1001 AGRDRGAGRRDRDAATDADGAGRRGGRKOKKDRKGGDDDKGGKKKKAATAKATK 942

QY 131 TTGGTATTCCTCAATTTTAGTAATTTTTCAGCTGAAGAAACAATCTCTGGTATTACAAAA 190
DB 941 WDDWDWDKWDGWDGAKDRKADDDGAGDKDDGKDDADDTDGTDKDDDDKDDKDDW 882

QY 191 AGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 250
DB 881 AKGTGWDATWAAATDWMWGWADADWWTWDAADADWADDDRDWADAWKWDADAWGART 822

QY 251 CATATGACCTACTTATGCAAGCAGTTTTCGAGGTTTATGCTATCATGATATGATG 310
DB 821 ADRRDRGDRAGRGSGARKRRDRKRDADDDKRDADDDAATTTTTRTDDTDKWKTD 762

QY 311 TTAATGACTTTAGGTAGAAATTTGAAGGTTCTTTATGAGAAATTTTGAACCTGAAAGCAAT 370
DB 761 TWRWAADRTWDRDDDDDRDRAGTAGRKWRRTWKRWRKRDRTWDDADDDTARDRRR 702

QY 371 GGTAACCTGAGAAATGCAAAAGCTACAAATTTTCTGCTGCTGCTGCTGCTGCTGCTG 430
DB 701 GDDGADAGKGTGRRRRDRATWDRDADWADAANWTTTDTDDTDKDRRRKRGARR 642

QY 431 GTGATAAAGTTTATAGTACTAGAGAAATAACGGCTGTTGACAGTCTCTTAATGTAA 490
DB 641 RRETARAAWDMWTWKAWDKWDKTRADRWDRWAADTTWTDARKADRWAKARWRARR 582

QY 491 ATGTTTGTATGATATGCTAGGTAGTATTCCTTTAGCACCTTATATGCTGCTGGTG 550
DB 581 DRARARADRRWTGKTTTATWTTWAARAAWAWAWATTTATWTTTWTWTTTWT 522

QY 551 TTGGTGCAGATTAATAAAGTTTATAGGTATATCATGCTGCTGCTGCTGCTGCTGCTG 610
DB 521 TTTTWTWTTAAWAAWATWATAWAAATAAATAAATAAATAAATAAATAAATAAATAA 462

QY 611 AGTTGGTGCACACTACCTCAAAATGTTAATATCTATGTTGTT 654
DB 461 TAWTWTWTTTWTWAAATTTTWTWTTWAAATTTATTTTWT 418
```

```
RESULT 2
CNS0100X 1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence sp6 end of BAC
BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL098379
VERSION AL098379.1 GI:5609990
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
```

d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBel0BAC11.

FEATURES

Location/Qualifiers

source 1..1101

organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACN03G04"

/clone\_lib="DrosBAC"

/plasmid="pBel0BAC11"

/note="end : sp6"

## ORIGIN

Query Match 7.5%; Score 60.8; DB 10; Length 1101;  
Best Local Similarity 17.6%; Pred. No. 0.0015;  
Matches 124; Conservative 275; Mismatches 306; Indels 1; Gaps 1;

QY 43 TTTGTTTTTCTTATCTGATGCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGA 102  
DB 389 KTTTTRTTKGGGTGKRTGNGTGTGAGKGGDKGDKGAKWTDTRTDDTKTDTWGKA 448

QY 103 GGACTTTATATAGGTAGTAAAGTTGATATCCCAATTTTAGTAATTTTTCAGCT 162  
DB 449 DTDWTDWTDWTTTWTWTTTGTGKWDTKTKWDDGTTWTDWTDWTDWTDWTDWTD 508

QY 163 GAAGAAACATTCCTGGTATTACAAAAGATTTTGGTGGTGGTGGTGGTGGTGGTGGT 222  
DB 509 DWADNAGWTDKDKTKAKKRTTDTKTKTGTGKTKDRGMRKTDTWGWDWRADRT 568

QY 223 ATAAATACTCACAGCAATTTTACAGCATATGACCTACTTATGCAAGCAAGTTTGA 282  
DB 569 DGARADRAANAKDDAGKADTDATKTDGWTGTTKADTTDDTDADWWDWWDKADATK 628

QY 283 GGGTTTAGTGGTATCATTTGGATATATGTTAATGACTTTTAGGGTGAAGTTTCAAGGTTCT 342  
DB 629 AGRKGGWGGKGGKWKTKTKKTKTCTGCTDTKGWDKDKWTDKDKRGGGDKDKDK 688

QY 343 TATGAGAAATTTGAACCTGAAAGCAATGGTACCTCGAGAAATAGCCAAAGCTTACAAATTT 402  
DB 689 GTGGDGDGSKAGDRAWAWRAKATRAAAADATAAATAAKATAWATKTTTCTTTTCTTT 748

QY 403 TTTGCTTTGCTCGAAATGCTACAATAGTATGATAATAGTTTATAGTACTAGAGATAAC 462  
DB 749 TTTTCTTTTCTTARKADDDDDKAKRDKDKGDDDDKDKAG-RDRDDDDTRKWK 807

QY 463 GCGTTTGTGCAAGTCTCTTAATGTAATGTTTCTTATGATATTTGCTAGTGTAGTATT 522  
DB 808 DRAGKKKAKKKHKTAWWADDTGTATWAKTATATKAKAKATKDBKRGGRGGGRGD 867

QY 523 CTTTATGACCTTATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 582  
DB 868 RGRGRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 927

QY 583 TCATTGCTCAAGTTTCTTATCAAGTAAAGTTTGGTGTCAACTCCCTCTTAATGTTAAT 642  
DB 928 DSKTWKTTWGTGDDTTTWTGRRNRGRGAGRRRGRKGGKAGGAGGKAGWAWADAWT 987

QY 643 ACTATGTTGTTTGGTGGGTTTATACCATAGGTTTGTAGGTGATAGGATGAGAGATGA 702  
DB 988 WTMDKDKWKKWTDGKGGGWDGKGGKAGKAGKRRKRRKRRKRRKRRKRRKRRKRRKRR 1047

QY 703 GAATAGCTTACCTCTACTGCAATATCTGACGTTCTCTAGAACTA 748  
DB 1048 KARWADDWDADYDDADADYWKWKAATWDBWCDCTAKWTANWKNOW 1093

## RESULT 3

EX173672/c  
LOCUS EX173672 829 bp DNA linear GSS 13-MAR-2003  
DEFINITION Dario rerio genomic clone DKEY-150M6, genomic survey sequence.  
ACCESSION EX173672

```
VERSION      BX173672.1  GI:28005377
KEYWORDS     GSS.
SOURCE       Danio rerio (zebrafish)
ORGANISM     Danio rerio
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
REFERENCE    1 (bases 1 to 829)
AUTHORS      Humphray, S.J., Huckle, E. and Durham, J.L.
TITLE        Direct Submission
JOURNAL      Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
              Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
              humphray@sanger.ac.uk Unpublished
COMMENT      This sequence was generated from the SP6 end of BAC 150M6. 150M6 is
              part of the Daniokey BAC Library created by R. Plasterk and N.V.
              Keygene. Further details:
              http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES     Location/Qualifiers
             source
               1..829
               /organism="Danio rerio"
               /mol_type="genomic DNA"
               /db_xref="taxon:7955"
               /clone="DKB1-150M6"
               /tissue_type="Testis"
               /note="vector pIndigoBAC-536"
ORIGIN
Query Match      7.0%; Score 56.8; DB 10; Length 829;
Best Local Similarity 43.9%; Pred. No. 0.013;
Matches 244; Conservative 0; Mismatches 312; Indels 0; Gaps 0;
Qy 60 T G A T G T G C T T T T T C T G A T C A A A T T T T T C T G A G G G A G A G A G A C T T T A T A G G T A G 119
Db 613 T G A T G A T G A T G A T G A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 554
Qy 120 T C A G T A T A A G T T G T T C C C A T T T T T A G T A A T T T T C A G T C G A G A A C A A T C T C G G 179
Db 553 T A A T A A T A A T G A T G A T G A T G A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 494
Qy 180 T A T T A C A A A A A G A T T T T G C G T T A G G T C T T G A T A A G C T C G A G A T A A T A C T C A C A G A A 239
Db 493 T A A T A G A A T A A T A A T A A T G A T G A T G A T G A T A A T A A T A A T A A T A A T A A T A A T A A 434
Qy 240 T T T T A C A G A T A T A G C C T A C T T A T G C A A G C A G T T T T C A G G G T T T A G T G G T A T C A T 299
Db 433 T A A T A A T A A T A A T A A T A A T A A A G A T A A T A A T A A T A A T A A T A A T A A T A A T A A 374
Qy 300 T G G A T A T T A T G T T A T A G A C T T T A G G G T A G A A T T T G A A G G T T C T T A T G A G A A T T T G A C C 359
Db 373 C A A T G A T G A T G A T G A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 314
Qy 360 T G A A A G A C A A T G G T A C C C T G A G A A T A G C A A A G C T A C A A A T T T T T T G C T T G T C T C G A A 419
Db 313 T A A T A A T A G T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 254
Qy 420 T G C T A C A A A T A G T G A T A A A G T T T A T A G T A C T A G A G A A T A A C G G C G T T G T T G C A A G T C 479
Db 253 T G A T G A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 194
Qy 480 T C T T A A T G T A A T G T T G T T A T A G A T A T T G C T A G T G G T A G T A T T C C T T T A G C A C C T T A T 539
Db 193 T G A T G A T G A T G A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 134
Qy 540 G T G T C G T G T G T G C G A G A T T A T A A A G T T T T T A G G T A T A T C A T C A T C C C T A A G T T T T C 599
Db 133 T A A T A C A A T A A T A A T A G A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A 74
Qy 600 T T A T C A A G T T A A G T T T 615
Db 73 A G A T G A C A C A C T T T T 58
RESULT 4
```





collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila melanogaster BAC library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial, *EcoRI* digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y<sup>+</sup>; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```

FEATURES
    source
    locust/qualifiers
    locust/qualifiers:11101
    organism="Drosophila melanogaster"
    map_type="genomic DNA"
    db_xref="tason:7227"
    clone="BACR48p13"
    clone_lib="RPCL-98"
    note="end . TERT"

```

ORIGIN	Query Match	Best Local Similarity	Best Local Similarity	Score 53;	DB 10;	Length 1101;
	Matches 123;	Conservative 182;	Mismatches 259;	Indels 2;	Gaps 1;	
Qy	125	ATAAAGTTGGTATTTCCCAATTTTAGTAATATTTTTCAGCTGAAGAACAATTCCTGGTATTA	184			
Db	476	AAAAAANAATAATTTTWTWTTTAAADDDKWAANWDTTTTAAWTWTWT	535			
Qy	185	CAAAAAAGATTTTTCGTTAGTCTTGATAGTCTGAGATAAATACTCACAGCAATTTTA	244			
Db	536	ATTAATATTTTWTWTKATKTDWTKAATAAATAAATTTTAAWTWTWT	595			
Qy	245	CACGATCATATGACCTACTTATGCAAGCAGTTTTCGAGGGTTTAGTGGTATCATATGAT	304			
Db	596	WTTAKTKWAKWAWNDATTTTWTWTTTWTWTTTAAATTTTWTWTTTAAWTWTWT	655			
Qy	305	ATTAATGTTAAATGACTTTTAGGGTGAATTTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAA	364			
Db	656	TTTAKTKTTTTTTTTTATTAATAAATAAATTTTWTWAAATTTTWTWTTTAAWTWTWT	715			
Qy	365	GACAATGGTACCTTGAGAATAGCCAAAGCTACAAATTTTGTCTGCTCAAAATGCTA	424			
Db	716	AKKWDGAKKATKTKKDKKAA--WAAADKKDKKGGKKGGKKGGKKGGKKKKKKGGK	773			
Qy	425	CAAAATAGTGATAATAAGTTTATAGTACTAGAGATAACGGCGTTGTGACAAAGTCTCTTA	484			
Db	774	GWGKKGGKAGDDBAKDKKKKKKAAATTTTKKKKKKKKKKAAKKKAAADRTTKTKRDA	833			
Qy	485	ATGTAATATGTTTGTATGATATGCTAGTGTAGTATTCCTTTTAGCACCTTATATGTGTG	544			
Db	834	AAAAAATAKTKDKGKKKKTKTKTTTKKKKKKKKKGGKKDAAAKKKKKGGTKKKKKGGK	893			
Qy	545	CTGGTGTGGTGACAGATATATAAGTTTATAGTATATCATGCTAAGTTTCTTATATC	604			
Db	894	GKGGGKGGKGGKDDDDAAAAAKKGGKKKKGGGGGGKKKKGGKKGGKKKKTKTKKK	953			
Qy	605	AAGTTAAGTTTGGTGTCACTACCTCTAAATGTTTAACTATATGTTGTGTGGTGGGGGT	664			
Db	954	KKKKKKDAAKGKTKKRAADAADAADWTDAATKKKKKKKKKKKKKKKKKAAAAATKKK	1013			
Qy	665	ATTACCATAAAGGTTGTAGGTGATAGG	690			
Db	1014	KKKKKKTKKKKKTKKK	1039			

LOCUS	CNS010DV	976 bp	DNA	linear	GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence sp6 end of BAC				

ACCESSION	BACN03N21 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION	AL098845
KEYWORDS	AL098845.1 GI:5610456
SOURCE	GSS.
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 976)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros_BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

FEATURES	source	Location/Qualifiers
		1..976
		/organism="Drosophila melanogaster"
		/mol_type="genomic DNA"
		/db_xref="taxon:7227"
		/clone="BACN0321"
		/clone_lib="DrosBAC"
		/plasmid="pBelobAC11"
		/note="end : Sp6"

Query Match	6.5%	Score 52.6;	DB 10;	Length 976;
Best Local Similarity	29.5%;	Pred. No. 0.13;		
Matches	94;	Conservative 92;	Mismatches 132;	Indels 1; Gaps 1;
Qy	16	TTTGTGTCAGTGTACGCTGAGTACAGTATTTGTTTTTTCTTATCTGATGCGTCTTTTCT	75	
Db	953	DWAKTKTKGKATWATRYDDRATATATGTWTTTWTWTATKRTGGTTTAT	894	
Qy	76	GATGCAAAATTTTCTGAAAGGAGGAGGAGCTTTATATAGTGTAGTCAGTATAAAGTTGGT	135	
Db	893	CKTGATTTGWTGDTGGWKGWTKDKWTAKRTTTATDWKATDKATAKAKKHGTGTTTAG	834	
Qy	136	ATTCCCAATTTTAGTA-ATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAAAGAT	194	
Db	833	WGDTRRRHAAKDGAWAWAWWWTADAGRWRTRAGNKTTGTTKTRMGTCATKRTDRGAK	774	
Qy	195	TTTTGCGCTTAGGTCCTGATAAGTCGAGATAAATACCTCACAGCAATTTTACACGATCATA	254	
Db	773	AKDVTGGKRGRTATRWATAWRTATTATATWTTAWTAWHTATTTTTTATTATTATRTD	714	
Qy	255	TGACCCCTACTATGTCGAAGCAGTTTTCAGCGGTTTAGTGGTATCATCTGGATATTATGTTAA	314	
Db	713	WGADTWWTATATKWTGTAKATCGAAAWTDTGWAITTTATTATWTTATHTTAKATWTT	654	
Qy	315	TGACTTTAGGGTGAATTT	333	
Db	653	ATWTATAWGTNNWAWTAT	635	

RESULT 10	630 bp	DNA	linear	GSS 28-JAN-2003
EX152436				
LOCUS	EX152436	Danio rerio genomic clone DREY-116F16,		
DEFINITION	Danio rerio			
ACCESSION	EX152436			
VERSION	EX152436.1	GI:27983930		
KEYWORDS	GSS.			
SOURCE	Danio rerio (zebrafish)			
ORGANISM	Danio rerio			



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 630)  
 AUTHORS Humphray, S.J., Huckle, E. and Durham, J.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished  
 COMMENT This sequence was generated from the SP6 end of BAC 116F16. 116F16 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: [http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/).

## FEATURES

source  
 1..630  
 Location/Qualifiers  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEY-116F16"  
 /tissue\_type="Testis"  
 /note="vector pIndigoBAC-536"

## ORIGIN

Query Match 6.3%; Score 51.2; DB 10; Length 630;  
 Best Local Similarity 48.9%; Pred. No. 0.26; Indels 0; Gaps 0;  
 Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;  
 QY 418 AATGCTACAAATAGTGAATAAAGTTATATAGTACTAGAGATAACGGGTTGTTGACAAG 477  
 DB 134 AATAATAATATAATAAATAAATAATATAATAATAATAATAATAATAATATATATTTATT 193  
 QY 478 TCTCTTAATGTAATGTTGTTGATATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTAT 537  
 DB 194 ATTATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 253  
 QY 538 ATGCTGCTGCTGTTGGTGCAGATTATATAAGTTTATAGTATATCATTCGCTCAAGTTT 597  
 DB 254 ATTGTTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 313  
 QY 598 TCTTATCAAGTTAAGTTTGGTGTCAACTACCTCTAAATGTTAATACTATCTATGTTGGT 657  
 DB 314 TATT 373  
 QY 658 GGGGGTTATTACCATAAGGTTGTAGGTGATGGCATGAGA 697  
 DB 374 ATTACTATTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 413

## RESULT 11

BX139987  
 LOCUS BX139987 844 bp DNA linear GSS 13-MAR-2003  
 DEFINITION Danio rerio genomic clone DKEY-99E7, genomic survey sequence.  
 ACCESSION BX139987  
 VERSION BX139987.1 GI:27971314

## KEYWORDS

## SOURCE

## ORGANISM

Danio rerio (zebrafish)  
 Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

## REFERENCE

AUTHORS Humphray, S.J., Huckle, E. and Durham, J.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished  
 COMMENT This sequence was generated from the T7 end of BAC 99E7. 99E7 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: [http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/).

## FEATURES

source  
 1..844  
 Location/Qualifiers  
 /organism="Danio rerio"

/mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEY-99E7"  
 /tissue\_type="Testis"  
 /note="vector pIndigoBAC-536"

## ORIGIN

Query Match 6.3%; Score 51; DB 10; Length 844;  
 Best Local Similarity 45.3%; Pred. No. 0.29; Indels 6; Gaps 1;  
 Matches 229; Conservative 0; Mismatches 270; Indels 6; Gaps 1;  
 QY 15 ATTGTTGTTAGGTTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTTTC 74  
 DB 304 AATTATATTGTCAGTCATTATTCCTTTTATTCTTTGATCATTAATAATAATAATAATGA 363  
 QY 75 TGATGCAAAATTTTCTGAAGGGAGGAGAGACTTTATATAGTAGGTAGTCAGTATATAAGTTGG 134  
 DB 364 TAATAATAATGATAATAATAATAATGATGATAATAATAATAATAATAATAATAATAATA 423  
 QY 135 TATTCCCAATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAAAGAT 194  
 DB 424 TAATAATAATAATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATGA 483  
 QY 195 TTTTGGCTTAGGTCCTTGATAAAGTCGAGATAAAATCTCACAGCAATTTTACCGATCAT 254  
 DB 484 TAATAATAATAATAATTTGATAATAATTTGATGATAATAATAATAATAATAATAATAATA 543  
 QY 255 TGACCCCTACTTATGCAAGCAGTTTTCGAGGGTTTAGTGGTATCATTCGGATATTATGTTAA 314  
 DB 544 TGATAATAATAATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 603  
 QY 315 TGACTTTAGGTTAGAAATTTGAAGTTCTTATGAGAAATTTTGAACCTGAAAGACATGGTA 374  
 DB 604 TGATGATA-----ATAATGATAATAATAACAATAATAATAATAATAATAATAATAATA 657  
 QY 375 CCCTGAGATAGCAAAAGCTACAAATTTTTCCTTGTCTCGAAATCTCTCAAAATAGTGA 434  
 DB 658 TAATGATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 717  
 QY 435 TAATAAGTTTATAGTACTAGAGATAACGGCGTTGTTGACAAGTCTCTTATGTAATGT 494  
 DB 718 CAATAATAATAATGACAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 777  
 QY 495 TTGTTATGATATTGCTAGTGGTAGT 519  
 DB 778 TACTAATAATAATGATAATAATAAT 802

## RESULT 12

CW973509 889 bp DNA linear GSS 21-DEC-2004  
 LOCUS A1AA-aaas9f06.b1 Ancylostoma caninum whole genome shotgun library  
 DEFINITION (A1AA-GSS 001) Ancylostoma caninum genomic, genomic survey sequence.  
 ACCESSION CW973509  
 VERSION CW973509.1 GI:56776238

## KEYWORDS

## SOURCE

## ORGANISM

Ancylostoma caninum (dog hookworm)  
 Ancylostoma caninum  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Ancylostomatidae; Ancylostomatinae; Ancylostomidae; Ancylostoma.  
 1 (bases 1 to 889)  
 Mitreva, M., McCarter, J.P., Pape, D., Ritter, E., Tsagareishvili, R., Ronko, I., Martin, J., Wylie, T., Danté, M., Meyer, R., Messina, D., Waterston, R.H., Clifton, S.W. and Wilson, R.  
 Genome Survey sequences from the parasitic nematode Ancylostoma caninum

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Unpublished (2004)  
 Contact: Mitreva M  
 Washington University in St. Louis  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

SOURCE ORGANISM	REFERENCE
<i>Drosophila melanogaster</i> (fruit fly)	
<i>Drosophila melanogaster</i>	
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
Ephydroidea; Drosophilidae; <i>Drosophila</i> .	
1 (bases 1 to 1094)	

REFERENCE	
AUTHORS	Repeyret, Endopocerygonia, Diphtheria, Brachymeria, Macromorphina,
TITLE	Ephydroidea; Drosophilidae; Drosophila.
JOURNAL	1 (bases 1 to 1094)
	Genoscope.
	Direct Submission
	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
	- Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.
FEATURES	Location/Qualifiers
source	1..1094

```
source
i. .1034
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN07B02"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"
```

```
ORIGIN
Query Match      6.2%; Score 50.4; DB 10; Length 1094;
Best Local Similarity 33.1%; Pred.No 0.42;
Matches 224; Conservative 93; Mismatches 355; Indels 4; Gaps 1;

QY    20   TTGTAGTGGTTACGCTCAGTACATTGTTTTTTTCCTATCTGATGGTGCTTTTTCGTATG 79
       |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db     1040 TTTTWTWTTTTTGGWTTTKTAKRWTTTTAKWTTTTTTTTTADWRRAA'TTTTWDBSWTAA 981
```

[illegible][illegible]

Qy	320	TAGGATGAGTATTGGAGGTCTTTTGACCTCGAAGCAAAATGGTACCCTC	
Db	744	WAAAAARATATTTTTTTTWWWWWTTTTTRTTTAAAAATAARAARTTTTTDGGGTG	685
Qy	380	AGAATAGCCAAAGCTACAAAATTTTTTGCTTCGAAATGCTACAAAATAGTGATAATA	439
Db	684	RTTWRATATAATTTAAKTTTTTTTGGRWARTTTTTTTRRAAAAWMTGWATGGRAVTTTG	625
Qy	440	AGTTTATAGTACTAGACAATAAACGGCGCTCTGTGACAAAGTCTCTTAATGTAAATCTTTGTT	499

624	RTTTGGKKAATKGTWAAATTTTTGGTTGGAATTTAAGWRBWTTTTTTAAATTTTTTTTTT	Db
500	ATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTCGTGTCGTGTCGTGCAG	Qy
564	TGRAAGTTTGGGGTGTTGCTGTGTTATKKGATTTTTTATKAAAWRYKGGGGAANK	Db
560	ATTATATAAAGTTTTTAGGTATATCATTCCTTAAGTTTTCTTATCAAGTTTAAGTTTGGT	Qy
504	TTAAATTTAAWTTTTTTTTTAAATWATKKTAAAGKAGGGTTTWTTTTTTTKTAACGGTGAAGGKT	Db



Qy	620	TCACCTACCCCTCAAACTGTTAACTACTATGTTTGGTGGGGGTTATTACCATAGGCTTG	679
Db	444	AATKTWMTTSSSTSGTTAAATAATGTTGTTTTTGGGGGAAAATGTTGGGGSTT	385
Qy	680	TAGGTGATAGGCATCA	695
Db	384	TTGGTTGKGGGGATAA	369

RESULT 14	CG807194	CG807194	471 bp	DNA	linear	GSS 10-NOV-2003
LOCUS		111807801.y1	1118	-	RescueMu Grid S	Zea mays genomic, genomic
DEFINITION						survey sequence.

AGNORUS	SOURCE	ORGANISM	REFERENCE
566.	Zea mays		1 (bases 1 to 471)
	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD	
		clade; Panicoideae; Andropogoneae; Zea.	

ORIGIN

Query Match	6.2%	Score 50.2	DB 10	Length 471;
Best Local Similarity	45.1%	Pred. No. 0.43;		
Matches 187;	Conservative	0;	Mismatches 228;	Indels 0;
Gaps 0;				

Qy 168 AACAAATTCCTGGTATTACAAAAAGATTTTTTCGCTTAGGTCTTGATAAGTCTGAGATAAA 227  
| | | | | | | | | | | | | | | |  
Db 61 TAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 120  
| | | | | | | | | | | | | | | |  
Qy 228 TACTCAGACGAATTTTACGCATCATATGACCCTACTTATCGAACAGATTTTTCAGGGTT 287

[illegible]

RESULT 15  
 BZ388480/c  
 LOCUS  
 DEFINITION BZ388480 670 bp DNA linear GSS 30-APR-2003  
 EINDL21TF EI\_10\_12 KB Entamoeba invadens genomic clone EINDL21,  
 genomic survey sequence.  
 ACCESSION BZ388480  
 VERSION BZ388480.1 GI:30235017  
 KEYWORDS GSS.  
 SOURCE Entamoeba invadens  
 ORGANISM Entamoeba invadens  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 670)  
 AUTHORS Wang, Z., Samuelson, J., Clark, C.G., Eichinger, D., Paul, J., van  
 Dellen, K., Hall, N., Anderson, I. and Loftus, B.  
 TITLE Gene discovery in the Entamoeba invadens genome  
 JOURNAL Mol. Biochem. Parasitol. 129 (1), 23-31 (2003)  
 PUBMED 12798503  
 COMMENT Other GSSs: EINDL21TR  
 Contact: Brendan Loftus  
 Department of Eukaryotic Genomics  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-3543  
 Fax: 301-838-0208  
 Email: enta@tigr.org  
 DNA was provided by Daniel Eichinger  
 Seq primer: TF  
 Class: sheared ends.

```

FEATURES
    source
        1. .670
            Location/Qualifiers
                /organism="Entamoeba invadens"
                /mol_type="genomic DNA"
                /strain="IP-1"
                /db_xref="taxon:33085"
                /clone="EINDL21"
                /clone_lib="EI 10 12 kb"
                /note="Vector: pHOS2; Site_1: BstXI; Total genomic DNA was isolated from early log phase trophozoites of E. invadens IP-1 using a Qiagen plant DNA extraction kit. A shotgun medium-size plasmid library (average insert size of 10 - 12 kb) was generated by random mechanical shearing of E. invadens genomic DNA, repairing the ends of DNA fragments with T4 Polymerase, adding BstXI adaptors and ligating into the BstXI site of a pUC-derived vector pHOS2."

```

ORIGIN

Query Match 6.2%; Score 50; DB 9; Length 670;  
Best Local Similarity 43.8%; Pred. No. 0.49;  
Matches 218; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

QY 117 TAGTCAGTATAAAGCTTGATATTCCTCCAAATTTTAGTAAATTTTTCAGCTCAAGAAACAATTC 176

Db 655 TAATA 606





Db 2806 TTATTGGAATTAAGTAATAGATAAATTTATTATTATTAGAAATTTTTTAAATTTTATAG 2865

QY 302 GATATTATGTTAATGACTTTAGGGTAGAATTTGAAAGTCTTATG-----AGAAATTTGAA 357

Db 2866 TAGTTATTATTTATATTTTGTGTAATTTATGATGTTTATGGGTTTATAGTTTGT 2925

QY 358 CCTGAAAGCAATGGGTACCTCGAATAGCCAAAGCTACAAATTTTGTGCTTGTCTCGA 417

Db 2926 TTTTATAGATATATATAAATGGAATATATAGGAATTTAGATTTTGTAGTTTGTGT 2985

QY 418 AATGCTACAAATAGTGAATAAGTTATAGTACTAGAGAAATACCGCGTGTGTGACAAG 477

Db 2986 TTTATATAGTAAATAGTGTAGATTTATTTATGTTAGTAGTTGTTTTTTTTTATG 3045

QY 478 TCTCTTAAGTAAATGTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTAT 537

Db 3046 TTTAGTAGTATTTTATGATGGTTTGTAGTTTATTTGTTTATAGTTGAAGGATAT 3105

QY 538 ATGTGTGCTGCTGTTGTTGTCAGATATATAAAAGTTTTTAGGTATATCATTTGCCCTAAGTTT 597

Db 3106 AGGATGTTTTTAGTTTTTGGATGTCGATPAAAGTTTTTATAATATTTTATATAGGAT 3165

QY 598 TCTTATCAAGTTAAGTTTGGTGT 620

Db 3166 TGTGTGAACGTAAGTTTTTATAT 3188

RESULT 2

US-10-829-826B-16

; Sequence 16, Application US/10829826B

; Publication No. US20050266397A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Blyn, Lawrence B.

; APPLICANT: Hall, Thomas A.

; APPLICANT: Massire, Christian

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES

; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829,826B

; CURRENT FILING DATE: 2004-04-22

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 20500

; TYPE: DNA

; ORGANISM: Coronavirus

US-10-829-826B-16

Query Match 5.6%; Score 45.6; DB 6; Length 20500;

Best Local Similarity 49.2%; Pred. No. 0.38;

Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 309 TGTAAATGACTTTAGGGTAGAATTTGAAAGTCTTATGAGAATTTTGAACCTGAAAGACA 368

Db 9378 TGGAACTGACCTAATGGGTGAATTTCTATGTTGTTATGATGAAGAGTTGCACAAAG 9437

QY 369 ATGGTACCTGAGAATAGCCAAAGCTACAAATTTTGTGTTGCTCGAAATGCTACAAA 428

Db 9438 AGTGCCACCAGATAATTTAGTTACTAACAAATTTAGTCATGGCTCTATGCGGCAATTAT 9497

QY 429 TAGTGATAATAAGTTTATAGTACTAGAGAATAACGGCTGTTGTGACAAAGTCTCTTAATGT 488

Db 9498 TAGTGTTAAGGAGAGTAGTTTCTCGCTGCCTAAATGTTGGAGAGTACTACTGTAGTGT 9557

QY 489 AAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG 548

Db 9558 TGATGATTATAATAAGTGGCTGGTGACAAATGTTTACACCAATTTTCTACTAGTACCGC 9617

QY 549 TGT 552

Db 9618 TATT 9621

RESULT 3

US-10-829-826B-14

; Sequence 14, Application US/10829826B

; Publication No. US20050266397A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Blyn, Lawrence B.

; APPLICANT: Hall, Thomas A.

; APPLICANT: Massire, Christian

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES

; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829,826B

; CURRENT FILING DATE: 2004-04-22

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14

; LENGTH: 27608

; TYPE: DNA

; ORGANISM: Coronavirus

US-10-829-826B-14

Query Match 5.6%; Score 45.6; DB 6; Length 27608;

Best Local Similarity 49.2%; Pred. No. 0.42;

Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 309 TGTAAATGACTTTAGGGTAGAATTTGAAAGTCTTATGAGAATTTTGAACCTGAAAGACA 368

Db 9378 TGGAACTGACCTAATGGGTGAATTTCTATGTTGTTATGATGAAGAGTTGCACAAAG 9437

QY 369 ATGGTACCTGAGAATAGCCAAAGCTACAAATTTTGTGTTGCTCGAAATGCTACAAA 428

Db 9438 AGTGCCACCAGATAATTTAGTTACTAACAAATTTAGTCATGGCTCTATGCGGCAATTAT 9497

QY 429 TAGTGATAATAAGTTTATAGTACTAGAGAATAACGGCTGTTGTGACAAAGTCTCTTAATGT 488

Db 9498 TAGTGTTAAGGAGAGTAGTTTCTCGCTGCCTAAATGTTGGAGAGTACTACTGTAGTGT 9557

QY 489 AAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG 548

Db 9558 TGATGATTATAATAAGTGGCTGGTGACAAATGTTTACACCAATTTTCTACTAGTACCGC 9617

QY 549 TGT 552

Db 9618 TATT 9621

RESULT 4

US-10-829-826B-15

; Sequence 15, Application US/10829826B

; Publication No. US20050266397A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Blyn, Lawrence B.

; APPLICANT: Hall, Thomas A.

; APPLICANT: Massire, Christian

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES

; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829,826B

; CURRENT FILING DATE: 2004-04-22

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 27608

; TYPE: DNA

; ORGANISM: Coronavirus

US-10-829-826B-15

Query Match 5.6%; Score 45.6; DB 6; Length 27608;  
Best Local Similarity 49.2%; Pred. No. 0.42;  
Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 309 TGTAAATGACCTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACA 368  
DB 9378 TGGAACTGACCTAATGGGTGAATTTCTATGTTGGTATGTTGATGAAGAGGTTGCACAAG 9437

QY 369 ATGTCACCTGAGAAATGACCAAGCTACAAATTTTGTCTGCTGAAATGCTACAAA 428  
DB 9438 AGTGCCACCAAGATAATTTAGTTACTAACAATATTGTAGCATGGCTCTATCGGCAATTAT 9497

QY 429 TAGTGATAATAAGTTTATGACTAGAGATAACGGGTTGTTGACAGTCTCTTAATGT 488  
DB 9498 TAGTGTAAAGAGAGTAGTTTCTCGCTGCCTAAATGGTTGGAGAGTACTACTGTTAGTGT 9557

QY 489 AAATGTTTGTATGATATGCTAGTGGTAGTATTCTTTAGCACCTTTATATGTTGCTGG 548  
DB 9558 TGATGATTATAAAGTGGGCTGGTGCAATGGTTTACACCAATTTTCTACTAGTACCGC 9617

QY 549 TGT 552  
DB 9618 TATT 9621

RESULT 5  
US-10-829-826B-17  
; Sequence 17, Application US/10829826B  
; Publication No. US20050266397A1  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; APPLICANT: Hofstadler, Steven A.  
; APPLICANT: Sampath, Rangarajan  
; APPLICANT: Blyn, Lawrence B.  
; APPLICANT: Hall, Thomas A.  
; APPLICANT: Massire, Christian  
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES  
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)  
; CURRENT APPLICATION NUMBER: US/10/829,826B  
; CURRENT FILING DATE: 2004-04-22  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 27635  
; TYPE: DNA  
; ORGANISM: Coronavirus  
; US-10-829-826B-17

Query Match 5.6%; Score 45.6; DB 6; Length 27635;  
Best Local Similarity 49.2%; Pred. No. 0.42;  
Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 309 TGTAAATGACCTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACA 368  
DB 9378 TGGAACTGACCTAATGGGTGAATTTCTATGTTGGTATGTTGATGAAGAGGTTGCACAAG 9437

QY 369 ATGTCACCTGAGAAATGACCAAGCTACAAATTTTGTCTGCTGAAATGCTACAAA 428  
DB 9438 AGTGCCACCAAGATAATTTAGTTACTAACAATATTGTAGCATGGCTCTATCGGCAATTAT 9497

QY 429 TAGTGATAATAAGTTTATGACTAGAGATAACGGGTTGTTGACAGTCTCTTAATGT 488  
DB 9498 TAGTGTAAAGAGAGTAGTTTCTCGCTGCCTAAATGGTTGGAGAGTACTACTGTTAGTGT 9557

QY 489 AAATGTTTGTATGATATGCTAGTGGTAGTATTCTTTAGCACCTTTATATGTTGCTGG 548  
DB 9558 TGATGATTATAAAGTGGGCTGGTGCAATGGTTTACACCAATTTTCTACTAGTACCGC 9617

QY 549 TGT 552  
DB 9618 TATT 9621

RESULT 6  
US-11-066-648A-1  
; Sequence 1, Application US/11066648A  
; Publication No. US20050260621A1  
; GENERAL INFORMATION:  
; APPLICANT: MCBRIDE, JERE W.  
; APPLICANT: WALKER, DAVID H.  
; APPLICANT: DOYLE, CHRISTOPHER KUYLER  
; TITLE OF INVENTION: AN IMMUNOREACTIVE 38-KDA FERRIC BINDING PROTEIN OF  
; FILE REFERENCE: CLPR:244US  
; CURRENT APPLICATION NUMBER: US/11/066,648A  
; CURRENT FILING DATE: 2005-02-25  
; PRIOR APPLICATION NUMBER: 60/547,499  
; PRIOR FILING DATE: 2004-02-25  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1044  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; US-11-066-648A-1

Query Match 5.6%; Score 45.2; DB 7; Length 1044;  
Best Local Similarity 47.8%; Pred. No. 0.19;  
Matches 131; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 82 AATTTTCTGAAGGAGGAGGAGGACTTTATATAGGTAGTCAGTATATAAGTTGTTATTCCTC 141  
DB 679 AGTTATTATTTTGGTAGAATTCGAGCTTCTGATAAGAAAAGTATCAAGCTGTAGTAAAA 738

QY 142 AATTTTAGTAATTTTTCAGCTGAAGAAAACAATTCCTGCTGTTATTAACAAAAAGATTTTTCG 201  
DB 739 AAGCTTGGTATTTTTCCTTAATCAAGAGACTACTGCTGACTATGATCAACATTAGTGGT 798

QY 202 TTAGGCTTGTAAAGTCTGAGATAAATCTACAGCAATTTTACACGATCATATGACCT 261  
DB 799 GGTGCTGTTACAAAAGCATGCAAAAATAAACAANAATGCTATAAATTTTGGAAATTCCTG 858

QY 262 ACTTATGCAAGCAGTATTTGTCAGGGTTTGTAGTGGTATCATTTGATATTTATGTTAATGACTTT 321  
DB 859 ACTAGTGAAGACACAGAAAAGTTTATGCTCAAGTTAATCAAGAATACCCCTATTGTAGAA 918

QY 322 AGGTGAGAAATTTGAAGGTTCTTATGAGAAATTTTG 355  
DB 919 GGTGTAAGAACTTTTCAAGAAAGTATTTAAAGACTTTTG 952

RESULT 7  
US-10-750-185-58930/c  
; Sequence 58930, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: PANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58930  
; LENGTH: 1773  
; TYPE: DNA  
; ORGANISM: Bovine  
; US-10-750-185-58930



```
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-22

Query Match      5.1%; Score 41.2; DB 6; Length 31028;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTCTACTGTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAAATGGTACCCTGAGATAGCCAAAGCTACAAATTTTGTGCTTT 410
Db 10509 TGCTCAGGTTGTCCAAATGGCCGTTCCAGGATTATACAACTCTGTTAAATTTGTAGCATG 10568

Qy 411 GTCTCGAAATGCTACAAATAGTACCTGAGATAGTTTATAGTACTAGAGAAATACCGCGTTGT 470
Db 10569 GCTTTATGCTGCTATCTTAATAATTTGTAATTTGGTTTGTACAAAGTGAAGTGTCTGT 10628

Qy 471 TGACAAGTCT 480
Db 10629 TGAAGATTTT 10638

RESULT 11
US-10-829-826B-26
; Sequence 26, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-26

Query Match      5.1%; Score 41.2; DB 6; Length 31028;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTCTACTGTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAAATGGTACCCTGAGATAGCCAAAGCTACAAATTTTGTGCTTT 410
Db 10509 TGCTCAGGTTGTCCAAATGGCCGTTCCAGGATTATACAACTCTGTTAAATTTGTAGCATG 10568

Qy 411 GTCTCGAAATGCTACAAATAGTACCTGAGATAGTTTATAGTACTAGAGAAATACCGCGTTGT 470
Db 10569 GCTTTATGCTGCTATCTTAATAATTTGTAATTTGGTTTGTACAAAGTGAAGTGTCTGT 10628

Qy 471 TGACAAGTCT 480
Db 10629 TGAAGATTTT 10638

RESULT 12
US-10-829-826B-23
; Sequence 23, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 31032
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-23

Query Match      5.1%; Score 41.2; DB 6; Length 31032;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTCTACTGTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAAATGGTACCCTGAGATAGCCAAAGCTACAAATTTTGTGCTTT 410
Db 10509 TGCTCAGGTTGTCCAAATGGCCGTTCCAGGATTATACAACTCTGTTAAATTTGTAGCATG 10568

Qy 411 GTCTCGAAATGCTACAAATAGTACCTGAGATAGTTTATAGTACTAGAGAAATACCGCGTTGT 470
Db 10569 GCTTTATGCTGCTATCTTAATAATTTGTAATTTGGTTTGTACAAAGTGAAGTGTCTGT 10628

Qy 471 TGACAAGTCT 480
Db 10629 TGAAGATTTT 10638

RESULT 13
US-10-829-826B-24
; Sequence 24, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 31100
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-24

Query Match      5.1%; Score 41.2; DB 6; Length 31100;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTCTACTGTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAAATGGTACCCTGAGATAGCCAAAGCTACAAATTTTGTGCTTT 410
```

```

Db 10509 TGCTCAGGTGTCCAAATGCGCGTTCAGGATTATATACAAATCGTTAAATTTTGTAGCATG 10568
Qy 411 GTCTCGAATGCTACAATAGTATTAATAGTTTATAGTACTAGAGAAATACGGCGTTGT 470
Db 10569 GCTTTATGCTGCTACTACTTAATAATTTGTAATGTTTGTACAAAGTGATAAGTGTCTGT 10628
Qy 471 TGACAAAGTCT 480
Db 10629 TGAAGATTTT 10638

```

```

RESULT 14
US-10-240-708-11
; Sequence 11, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/240,708
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-11

```

```

Query Match 4.9%; Score 40; DB 6; Length 6317;
Best Local Similarity 44.8%; Pred. No. 4.9; Indels 0; Gaps 0;
Matches 154; Conservative 0; Mismatches 190;

Qy 12 AAGATTGTTGTAGGTGTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTT 71
Db 1588 AAGTTTTTTTATTTTATGAAATATTTTAAATATATAGAGAAATGAAAGAGTAT 1647

Qy 72 TTCTGATGCAAAATTTTCTGAAGGGAGGAGAGCTTTTATATAGGTAGTCAGTATAAAGT 131
Db 1648 TGTTAATAAATATTAATGAATAGATTTAATAATTTGTTAATATTTTGTGCGTATGTAT 1707

Qy 132 TGGTATTTCCAAATTTAGTAAATTTTTCAGCTGAAGAAACAAATTCCTGGTATTACAAA 191
Db 1708 TAATATTTTAAATTTTATGATTAATAATTTTAAAGTATTTTATAATATTATAAGATA 1767

Qy 192 GATTTTTGCGTTAGGTCTTGATAAGTCTGAGATAAATACTCACAGCAATTTTACACGATC 251
Db 1768 TTTTATTTTAAATATTTTATGATATGATATTTTAAAGTAAGGATAAATATATAGATAT 1827

Qy 252 ATATGACCCCTACTTATGCAAGAGAGTTTTCAGGGTTTAGTGGTATCATTGGATATTATGT 311
Db 1828 AATTATATTATAATTTTAAAGAAAATTTTATAATTTTTTAAATATTATTTGAAAATTAAT 1887

Qy 312 TAATGACTTTTGGGTAGAAATTTTCAGGTTCTTATGAGAAATTTTG 355
Db 1888 TTTTATTTTATTTATTTGTTTAAATGTTTATAGTTATTTTG 1931

```

```

RESULT 15
US-10-750-185-39647
; Sequence 39647, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 39647
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Bovine 19866880742947
US-10-750-185-39647

Query Match 4.8%; Score 39.2; DB 6; Length 1289;
Best Local Similarity 48.6%; Pred. No. 4.6; Indels 0; Gaps 0;
Matches 107; Conservative 0; Mismatches 113;

Qy 295 ATCATTGGATATTTATGTTAATGACCTTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTT 354
Db 193 ATTATTTTGGTTTATTTTAAACAGGATGAATTAACCTTGTCTGATTTCTGAAGTGGATAAC 252

Qy 355 GAACCTGAAAGACAATGGTACCCCTGAGAAATAGCCAAAGCTACAAATTTTTTGTCTTTGTCT 414
Db 253 CAAAAGCGAGGGAACCGGCATTTATGAAGACAAGCAGAAAGACACTTGGATACCTTTAAAT 312

Qy 415 CGAAATGCTACAAATAGTGAATAATAGTTTATAGTACTAGAGAAATAACGGCTTCTTGAC 474
Db 313 AAAAAAGAACCGAGAACTAGATATGAAGAAAAGAGCTAGAGGTTTATCCATGATTTTATA 372

Qy 475 AAGTCTCTTAATGTTAAATGTTTGTGTTATGATATTGCTAGTG 514
Db 373 AAACCTTTTATTACACCTTAGTCATGATTTATTAAGTG 412

```

Search completed: December 28, 2005, 15:34:25  
Job time : 842 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 12:12:19 ; Search time 868 Seconds  
(without alignments)  
7745.397 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaaagattgt.....ttggattgattgctgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813	100.0	813	3	US-09-811-007-45
2	813	100.0	813	5	US-10-062-624-45
3	813	100.0	813	5	US-10-062-051-45
4	813	100.0	813	5	US-10-062-920-45
5	813	100.0	813	5	US-10-062-349-45
6	162.2	20.0	840	5	US-10-731-554-45
7	162.2	20.0	840	5	US-10-314-639-23
8	162.2	20.0	840	5	US-10-901-714-23
9	162.2	20.0	840	8	US-10-901-714-23
10	162.2	20.0	840	8	US-10-901-774-23
11	150.8	18.5	852	5	US-10-059-964-25
12	150.8	18.5	852	5	US-10-314-639-25
13	150.8	18.5	852	8	US-10-901-714-25
14	150.8	18.5	852	8	US-10-901-774-25
15	147.2	18.1	840	3	US-09-811-007-41
16	147.2	18.1	840	5	US-10-062-624-41
17	147.2	18.1	840	5	US-10-062-051-41
18	147.2	18.1	840	5	US-10-062-920-41
19	147.2	18.1	840	8	US-10-680-349-41
20	147.2	18.1	840	8	US-10-731-554-41
21	147.2	18.1	843	5	US-10-059-964-47
22	147.2	18.1	843	5	US-10-314-639-47
23	147.2	18.1	843	8	US-10-901-714-47

24	147.2	18.1	843	8	US-10-901-774-47	Sequence 47, Appl
25	146.6	18.0	726	5	US-10-059-964-61	Sequence 61, Appl
26	146.6	18.0	726	5	US-10-314-639-61	Sequence 61, Appl
27	146.6	18.0	726	5	US-10-901-714-61	Sequence 61, Appl
28	146.6	18.0	726	5	US-10-901-774-61	Sequence 61, Appl
29	146.4	18.0	849	9	US-10-138-162-48	Sequence 48, Appl
30	135	16.6	828	5	US-10-059-964-27	Sequence 27, Appl
31	135	16.6	828	5	US-10-314-639-27	Sequence 27, Appl
32	135	16.6	828	8	US-10-901-714-27	Sequence 27, Appl
33	135	16.6	828	8	US-10-901-774-27	Sequence 27, Appl
34	133.8	16.5	840	5	US-10-059-964-59	Sequence 59, Appl
35	133.8	16.5	840	5	US-10-314-639-59	Sequence 59, Appl
36	133.8	16.5	840	8	US-10-901-714-59	Sequence 59, Appl
37	133.8	16.5	840	8	US-10-901-774-59	Sequence 59, Appl
38	129.6	15.9	843	5	US-10-059-964-5	Sequence 5, Appl
39	129.6	15.9	843	5	US-10-314-639-5	Sequence 5, Appl
40	129.6	15.9	843	8	US-10-901-714-5	Sequence 5, Appl
41	129.6	15.9	843	8	US-10-901-774-5	Sequence 5, Appl
42	126.4	15.5	837	5	US-10-062-994-9	Sequence 9, Appl
43	126.4	15.5	837	5	US-10-059-964-9	Sequence 9, Appl
44	126.4	15.5	837	5	US-10-062-994-9	Sequence 9, Appl
45	126.4	15.5	837	5	US-10-314-639-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-09-811-007-45  
; Sequence 45, Application US/09811007  
; Publication No. US20030185849A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152C1F2  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 45  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9  
US-09-811-007-45

Query Match	100.0%	Score 813;	DB 3;	Length 813;
Best Local Similarity	100.0%	Pred. No. 9.2e-181;		
Matches 813;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAATTACAAAAGATTGTTGAGGTGTTACGCTGAGTACATTGTTTCTTATCT	60	
Db	1	ATGAATTACAAAAGATTGTTGAGGTGTTACGCTGAGTACATTGTTTCTTATCT	60	
Qy	61	GATGGTCTTTTCTGATGCAATTTTCTGAAGGGAGGAGGACCTTTATAGGTAGT	120	
Db	61	GATGGTCTTTTCTGATGCAATTTTCTGAAGGGAGGAGGACCTTTATAGGTAGT	120	
Qy	121	CAGTATAAAGTGGTATTCCCAATTTTGTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGT	180	
Db	121	CAGTATAAAGTGGTATTCCCAATTTTGTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGT	180	
Qy	181	ATTACAAAAAGATTTTTCGGTGTAGGTCTTTGATAAGTCTGAGATAAATCTCAGCAAT	240	
Db	181	ATTACAAAAAGATTTTTCGGTGTAGGTCTTTGATAAGTCTGAGATAAATCTCAGCAAT	240	
Qy	241	TTTACAGATCATATGACCTTACTTATGCAAGCAATTTTCAGCGTTTGTAGGTATCAT	300	

Db 241 TTTACACGATCATATGACCCCTACTTATGCAAGCAGTTTTTGCAGGGTTTAGTGGTATCAATT 300  
Qy 301 GGTATATTGTTAAATGACTTTAGGTAGATTTGAGGTTCTTATCAGAAATTTTGAACCT 360  
Db 301 GGTATATTGTTAAATGACTTTAGGTAGATTTGAGGTTCTTATCAGAAATTTTGAACCT 360  
Qy 361 GAAAGACAAATGGTACCCCTGAGATAGCCAAAGCTACAAATTTTTTTCCTTTGCTCGAAAT 420  
Db 361 GAAAGACAAATGGTACCCCTGAGATAGCCAAAGCTACAAATTTTTTTCCTTTGCTCGAAAT 420  
Qy 421 GCTACAAATAGTGATTAATAAGTTTATAGTACTAGAGAAATAACCGCGTTGTTGACAAAGTCT 480  
Db 421 GCTACAAATAGTGATTAATAAGTTTATAGTACTAGAGAAATAACCGCGTTGTTGACAAAGTCT 480  
Qy 481 CTTAAATGTAATCTTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
Db 481 CTTAAATGTAATCTTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
Qy 541 TGTGCTGGTGTGGTGACGATTAATAAAGTTTATAGTATATCATTTGCTTAAGTTTCT 600  
Db 541 TGTGCTGGTGTGGTGACGATTAATAAAGTTTATAGTATATCATTTGCTTAAGTTTCT 600  
Qy 601 TATCAAGTTTAAAGTTTGGTGCAACTACCCCTCTAAATGTTTAAATGTTTGGTGGG 660  
Db 601 TATCAAGTTTAAAGTTTGGTGCAACTACCCCTCTAAATGTTTAAATGTTTGGTGGG 660  
Qy 661 GGTATATACCAATAGGTTGAGTGATAGGCATGAGAGAGTAGAATAGCTTACCACCT 720  
Db 661 GGTATATACCAATAGGTTGAGTGATAGGCATGAGAGAGTAGAATAGCTTACCACCT 720  
Qy 721 ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGTCTTCTGCTACTTTAAATAGTATAT 780  
Db 721 ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGTCTTCTGCTACTTTAAATAGTATAT 780  
Qy 781 TTTGGTTGGAGATTGGATTTAGATTGGCGTA 813  
Db 781 TTTGGTTGGAGATTGGATTTAGATTGGCGTA 813

RESULT 2

US-10-062-624-45  
; Sequence 45, Application US/10062624  
; Publication No. US20020115840A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2/D1  
; CURRENT APPLICATION NUMBER: US/10/062,624  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 45  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9  
US-10-062-624-45

Query Match 100.0%; Score 813; DB 5; Length 813;  
Best Local Similarity 100.0%; Pred. No. 9.2e-181;  
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAATTTACAAAAGATTGTTGAGGTGTTACGCTCAGTACATTTCTGTTTTTCTTATCT 60  
Db 1 ATGAATTTACAAAAGATTGTTGAGGTGTTACGCTCAGTACATTTCTGTTTTTCTTATCT 60  
Qy 61 GATGGTCTTTTCTGATGCAAAATTTTCTGAAAGGAGGAGGACTTTATATAGTACT 120

Db 61 GATGGTCTTTTCTGATGCAAAATTTTCTGAAAGGAGGAGGACTTTATATAGTACT 120  
Qy 121 CAGTATAAGTTGGTATTTCCCAATTTTATAGTAAATTTTTCAGCTGCAAGAACAAATTTCTGGT 180  
Db 121 CAGTATAAGTTGGTATTTCCCAATTTTATAGTAAATTTTTCAGCTGCAAGAACAAATTTCTGGT 180  
Qy 181 ATTACAAAAGATTTTTCGCTTAGCTCTTGATAAGTCTGAGATAAATACTCACAGCAAT 240  
Db 181 ATTACAAAAGATTTTTCGCTTAGCTCTTGATAAGTCTGAGATAAATACTCACAGCAAT 240  
Qy 241 TTTACACGATCATATGACCCCTACTTATGCAAGCAGCTTTTTCAGGCTTTAGTGGTATCAT 300  
Db 241 TTTACACGATCATATGACCCCTACTTATGCAAGCAGCTTTTTCAGGCTTTAGTGGTATCAT 300  
Qy 301 GGATATATGTTAAAGTCTTTAGGCTAGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCT 360  
Db 301 GGATATATGTTAAAGTCTTTAGGCTAGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCT 360  
Qy 361 GAAAGACAAATGCTACCCCTGAGAAATAGCCAAAGCTACAAATTTTTTTCCTCGAAAT 420  
Db 361 GAAAGACAAATGCTACCCCTGAGAAATAGCCAAAGCTACAAATTTTTTTCCTCGAAAT 420  
Qy 421 GCTACAAATAGTGATTAATAAGTTTATAGTACTAGAGAAATAACCGCGTTGTTGACAAAGTCT 480  
Db 421 GCTACAAATAGTGATTAATAAGTTTATAGTACTAGAGAAATAACCGCGTTGTTGACAAAGTCT 480  
Qy 481 CTTAAATGTAATCTTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
Db 481 CTTAAATGTAATCTTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540  
Qy 541 TGTGCTGGTGTGGTGACGATTAATAAAGTTTATAGTATATCATTTGCTTAAGTTTCT 600  
Db 541 TGTGCTGGTGTGGTGACGATTAATAAAGTTTATAGTATATCATTTGCTTAAGTTTCT 600  
Qy 601 TATCAAGTTTAAAGTTTGGTGTCAACTACCCCTCTAAATGTTTAAATGTTTGGTGGG 660  
Db 601 TATCAAGTTTAAAGTTTGGTGTCAACTACCCCTCTAAATGTTTAAATGTTTGGTGGG 660  
Qy 661 GGTATATACCAATAGGTTGAGTGATAGGCATGAGAGTAGAATAGCTTACCACCT 720  
Db 661 GGTATATACCAATAGGTTGAGTGATAGGCATGAGAGTAGAATAGCTTACCACCT 720  
Qy 721 ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGTCTTCTGCTACTTTAAATAGTATAT 780  
Db 721 ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGTCTTCTGCTACTTTAAATAGTATAT 780  
Qy 781 TTTGGTTGGAGATTGGATTTAGATTGGCGTA 813  
Db 781 TTTGGTTGGAGATTGGATTTAGATTGGCGTA 813

RESULT 3

US-10-062-051-45  
; Sequence 45, Application US/10062051  
; Publication No. US20030073095A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/10/062,051  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US/09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 45  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis



```
QY 781 TTTGGTTGGAGATTGGATTAGATTGGCTA 813
Db |||||
QY 781 TTTGGTTGGAGATTGGATTAGATTGGCTA 813
Db |||||

RESULT 5
US-10-680-349-45
; Sequence 45, Application US/10680349
; Publication No. US20040198951A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/680,349
; CURRENT FILING DATE: 2003-10-07
; PRIOR APPLICATION NUMBER: US/10/062,624
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-10-680-349-45

Query Match 100.0%; Score 813; DB 8; Length 813;
Best Local Similarity 100.0%; Pred. No. 9.2e-181;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTACAAAAGATTGTTGAGGTGTTACCTCAGTACATTTGTTTTTCTTATCT 60
Db |||||
QY 1 ATGAATTACAAAAGATTGTTGAGGTGTTACCTCAGTACATTTGTTTTTCTTATCT 60
Db |||||

QY 61 GATGGTGTCTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATAGTAGT 120
Db |||||
QY 61 GATGGTGTCTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATAGTAGT 120
Db |||||

QY 121 CAGTATAAAGTTGGTATCCCAATTTTCTGAGTAAATTTTTCAGCTGAAGAACTTCCTGT 180
Db |||||
QY 121 CAGTATAAAGTTGGTATCCCAATTTTCTGAGTAAATTTTTCAGCTGAAGAACTTCCTGT 180
Db |||||

QY 181 ATTACAAAAGATTGTTGCGTGTAGTCTTGTAGTAAGTCTGAGATAAATACTCACAGCAAT 240
Db |||||
QY 181 ATTACAAAAGATTGTTGCGTGTAGTCTTGTAGTAAGTCTGAGATAAATACTCACAGCAAT 240
Db |||||

QY 241 TTTACACGATCATATGACCTTACTTATGCAAGCACTTTTTCAGGGTGTAGTGTATCAT 300
Db |||||
QY 241 TTTACACGATCATATGACCTTACTTATGCAAGCACTTTTTCAGGGTGTAGTGTATCAT 300
Db |||||

QY 301 GGATATTAATGTTAAATGACTTTTAGGGTAGAATTTTGAAGGTTCTTTATGAGAACTTGAACCT 360
Db |||||
QY 301 GGATATTAATGTTAAATGACTTTTAGGGTAGAATTTTGAAGGTTCTTTATGAGAACTTGAACCT 360
Db |||||

QY 361 GAAAGACAATGGTACCTCGAGATAGCCAAAGCTACAAATTTTTCGTTTGTCTCGAAAT 420
Db |||||
QY 361 GAAAGACAATGGTACCTCGAGATAGCCAAAGCTACAAATTTTTCGTTTGTCTCGAAT 420
Db |||||

QY 421 GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATACCGCGTGTGTGACAGTCT 480
Db |||||
QY 421 GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATACCGCGTGTGTGACAGTCT 480
Db |||||

QY 481 CTTAAATGTAATGTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCAGCTTATAG 540
Db |||||
QY 481 CTTAAATGTAATGTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCAGCTTATAG 540
Db |||||

QY 541 TGTGCTGGTGTGGTGCAGATTATATAAAGTTTTTAGGTATATCAATGTCCTAAGTTTCT 600
Db |||||
```

QY 361 GAAAGCAATGGTACCTGAGATAGCCAAAGCTACAAATTTTGTGCTCGAAAT 420  
 DB 361 GAAAGCAATGGTACCTGAGATAGCCAAAGCTACAAATTTTGTGCTCGAAAT 420  
 QY 421 GCTACAAATAGTGAATAGTTTATAGTACTAGAGATAACCGGTTGTTGCAAGTCT 480  
 DB 421 GCTACAAATAGTGAATAGTTTATAGTACTAGAGATAACCGGTTGTTGCAAGTCT 480  
 QY 481 CTTAAATGTAATGTTGTTATGATATTGCTAGTGTAGTATTCCTTTAGCACCTTATATG 540  
 DB 481 CTTAAATGTAATGTTGTTATGATATTGCTAGTGTAGTATTCCTTTAGCACCTTATATG 540  
 QY 541 TGTGCTGGTGTGTCAGATATATATAAGTTTATAGTATATATGCTTACCTTACCTAGTTTCT 600  
 DB 541 TGTGCTGGTGTGTCAGATATATATAAGTTTATAGTATATATGCTTACCTTACCTAGTTTCT 600  
 QY 601 TATCAAGTTAAGTTTGTGTCACCTCCCTCTAAATGTTTAACTATGTTTGTGTTGGG 660  
 DB 601 TATCAAGTTAAGTTTGTGTCACCTCCCTCTAAATGTTTAACTATGTTTGTGTTGGG 660  
 QY 661 GGTATTATACCAATAGTTTGTAGTGTAGGATGATGAGAGATAGAAATAGCTTACCATCCT 720  
 DB 661 GGTATTATACCAATAGTTTGTAGTGTAGGATGATGAGAGATAGAAATAGCTTACCATCCT 720  
 QY 721 ACTGCATATCTGAGTTCCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATAGTATTAT 780  
 DB 721 ACTGCATATCTGAGTTCCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATAGTATTAT 780  
 QY 781 TTTGGTTGGGAGATTGGATTTCGCGCTA 813  
 DB 781 TTTGGTTGGGAGATTGGATTTCGCGCTA 813

RESULT 7

US-10-059-964-23  
 ; Sequence 23, Application US/10059964  
 ; Publication No. US20020120115A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rikihisa, Yasuko  
 ; APPLICANT: Ohasi, No. US20020120115A1io  
 ; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
 ; TITLE OF INVENTION: Chaffeensis  
 ; FILE REFERENCE: 22727/04021  
 ; CURRENT APPLICATION NUMBER: US/10/059,964  
 ; CURRENT FILING DATE: 2002-01-28  
 ; EARLIER APPLICATION NUMBER: 09/314,701  
 ; EARLIER FILING DATE: 1999-05-19  
 ; NUMBER OF SEQ ID NOS: 66  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 23  
 ; LENGTH: 840  
 ; TYPE: DNA  
 ; ORGANISM: Ehrlichia chaffeensis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(840)  
 US-10-059-964-23

Query Match 20.0%; Score 162.2; DB 5; Length 840;  
 Best Local Similarity 53.8%; Pred. No. 9.4e-28;  
 Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGTGTGTAGCTGAGTACATTTGTTTTTCTTATCTGATGCTGCTTTTCTGTGCA 81  
 DB 34 GTACTTGATCTCTATTATCTTCTTATCTTATGAACTTTTTCAGCTATAAATCAAT 93  
 QY 82 AATTTTCTGAAGGGAGAGAGGACTTTTATAGGTAGTCAAGTATAAAGTTGGTATCCC 141  
 DB 94 CATACAGGAAATACACACTAGTGGTATATATATACAGGGCAGTATAGACAGGATATCC 153  
 QY 142 AATTTTAGTAAATTTTACGTGAGAAACAATTCCTGTTATACAAAAGATTTTTCG 201

DB 154 CATTTTAGCAATTTCTCAGTAAAGAACTAATGTTGTATACAACTACCACTAGTAGGATAT 213  
 QY 202 TTAGGTCTTGTAGTCT---GAGATAAATCTACAGCAATTTTACAGCATCATATGAC 258  
 DB 214 AAAAAAGTGGCTCTTCTATCGATCTTAACACTTATTCAAGCTTCCATATACT 273  
 QY 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTCTAGTGGTATCATTTGGATATTATGTTAATGAC 318  
 DB 274 GTTACATTTCAAGATAATGCTCTAGTTTCAGTGGAGCAATTTGGATATTCTTACCCGMA 333  
 QY 319 TTT---AGGGTAGAATTTGAAGTTTCTTATAGAAATTTTGAACCTGAAAGACAATGGTAC 375  
 DB 334 AGCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAAATTTGATGTCAAAGATCCTAAAGAC 393  
 QY 376 CTTGAGAAATAGCCAAAGCTACAAATTTTGTCTCTCGAAATCTCAAAATAGTATGAT 435  
 DB 394 TACTCAGCAAAAAGATGCTTTTATAGTTTGTCTAGCAGCAATTAATACGCTACTACTGTT 453  
 QY 436 AAT-----AAGTTTATAGTACTAGAGATAACCGGTTGTTGCACAAGTCTCTTAAT 486  
 DB 454 CTTGATGCTCAAAAATATACAGTTATGAAGAAATATGGCTTATCTGTTGCATCAATCATG 513  
 QY 487 GTAAATGTTTGTATGATATGCTAGTGGTGTAGTATTCTTTAGCACCTTATATGTTGCT 546  
 DB 514 ATCAATGGTTGTTATGATCTATCTTTTAAATTTTAGTCGTATCACCTTATATATGTCGA 573  
 QY 547 GGTGTTGGTGCAGATTATATAAGTTTTTAGGTATATCATTCGCTTAAGTTTCTTATCAA 606  
 DB 574 GGTATTGGTGAAGATTTCATTGAATTTTGTGATCTTTGCACATTTAAACTTGCCTTATCAA 633  
 QY 607 GTTAGTTTGGTGTCAACTACCTCTAAATGTTTAACTATGTTGTTGGTGGGGTTAT 666  
 DB 634 GGAATACTAGTATTATGTTTATCTTCTTCTTAAGATTAAATGTTTGGTGGGGTAC 693  
 QY 667 TACCATAAGTTTGTAGTGTAGGATGAGAGATAGAAATAGCTTACCATCTACTGCA 726  
 DB 694 TATCATAGATTATAGGAAATTAATTTTAAATTT---AAATGTTAAACCATGTTGTACA 750  
 QY 727 TTATCTGAGCTTCTAGAACTACTTTCAGCTTCTGCTACTTCTTAAATACTGATTTTGGT 786  
 DB 751 CTTGATGAATTTCTTAAAGCAACTTCTGCAGTAGCTACACTTAATGTTGCTTATTTGGT 810  
 QY 787 TGGGAGATTGGATTAGATTTC 809  
 DB 811 GGTGAAGCTGGAGTAAAGTTTAC 833

RESULT 8

US-10-314-639-23  
 ; Sequence 23, Application US/10314639  
 ; Publication No. US20030103991A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rikihisa, Yasuko  
 ; APPLICANT: Ohasi, No. US20030103991A1io  
 ; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
 ; TITLE OF INVENTION: Chaffeensis  
 ; FILE REFERENCE: 22727/04021  
 ; CURRENT APPLICATION NUMBER: US/10/314,639  
 ; CURRENT FILING DATE: 2002-12-09  
 ; PRIOR APPLICATION NUMBER: US/09/314,701  
 ; PRIOR FILING DATE: 1999-05-19  
 ; NUMBER OF SEQ ID NOS: 66  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 23  
 ; LENGTH: 840  
 ; TYPE: DNA  
 ; ORGANISM: Ehrlichia chaffeensis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(840)  
 US-10-314-639-23

Query Match

20.0%; Score 162.2; DB 5; Length 840;

```
Best Local Similarity 53.8%; Pred. No. 9.4e-28;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 GTACTTGCATCTCTATTATCATTTCTTATCTATCTGAACTCCTTTTCAGCTATAAATCATAAT 93

QY 82 AATTTTTCTGAAGGAGGAGACTTTATATAGGTAGTCAGTATAAAGTTGGTATCC 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 CATACAGGAATAAACAATGTGTATATATATACAGGCGAGTATAGACAGGAGTATCC 153

QY 142 AATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAGATTTTGGG 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 CATTTTAGCAATTTCTCAGTAAAGAACTAATGTTGATACAAATACAACTAGTAGATAT 213

QY 202 TTAGGCTTGTATAGTCT---GAGATAAATCTCACAGCAATTTTACAGATCATATGAC 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 AAAAAAAGTCGCTCTTCTATCGATCCTAACACTTATTCAAACCTTTCAAGGTCATATACT 273

QY 259 CCTACTTATGCAAGCAGTTTTCAGAGGTTTAGTGGTATCATTTGGATATTATGTTAATGAC 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 GTTCAATTCAGATTAATGCTGCTAGTTTCAGTGGAGCAATGGATATTCTTACCCGAA 333

QY 319 TTT---AGGTAAGATTTGAAGTCTTCTATGAGATTTTGAACCTGAAAGACAATGGTAC 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 AGCTAAGACTTGAAGTCTTACGAAAAATTTGATGTCAAAGATCCTAAGAC 393

QY 376 CCGTGAATAGCAAGCTACAAATTTTTCGTTCTGCGAAATGCTCAAAATAGTAT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCACGTAATACGCTACTACTGTT 453

QY 436 AAT-----AGTTTATAGTACTAGAGATAACGGGTTGTTGACAAAGTCTCTTAAT 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 CCTGATGCTCAAAAATATACAGTTATGAAGAAATAATGGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTTGTATGATATTGCTAGTGTAGTATTCCTTTAGCACCTTATATGTGTCT 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 ATCAATGGTTGTATGATCTATCTTTTAAATTTAGTCGTATCACCTTATATATGTGCA 573

QY 547 GGTGTTGGTCAGATATATAAAGTTTGTAGGTATATCATTTGCCCTAAAGTTTCTTATCAA 606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 574 GGTATTTGGTGAAGATTTCAATTGAATTTTGTGATCTTTTGCACATTTAAACCTTGTCTTAT 633

QY 607 GTTAAGTTTGGTCTCAACTACCTCTAAATGTTTAACTATGTTGTTGGTGGGTTAT 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 634 GGAAACTAGTATTAGTTATTACTTCTTTCTTAAGATTAAATGTTGCTGGTGGGTAC 693

QY 667 TACCATAAGGTTGTAGGTGATAGGCATGAGAGATAGAAATAGCTTTACCATCCTACTGCA 726
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 694 TATCATAGATTATAGGGAATAAATTTAAATAATTT---AAATGTTAAACCATGTTGTACA 750

QY 727 TTATCTGACGTTCTTAGAATCTCTCAGCTTCGTCTACTTTTAAATCTGATATTATTTGGT 786
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 CTTGATGAATTTCTTAAAGCAACTTCTGCAAGTACACATTAATGTTGCTTATTTTGGT 810

QY 787 TGGGAGATTGGATTTAGATTTGC 809
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 GGTGAAGCTGGAGTAAAGTTTAC 833
```

RESULT 9

```
US-10-901-714-23
; Sequence 23, Application US/10901714
; Publication No. US2004026533A1
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia
; TITLE OF INVENTION: CHAFFENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/901,714
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 09/314,701
```

```
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; US-10-901-714-23

Query Match 20.0%; Score 162.2; DB 8; Length 840;
Best Local Similarity 53.8%; Pred. No. 9.4e-28;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 GTACTTGCATCTCTATTATCATTTCTTATCTATCTGAACTCCTTTTCAGCTATAAATCATAAT 93

QY 82 AATTTTTCTGAAGGAGGAGGAGCACTTTATATAGGTAGTCAGTATAAAGTTGGTATCC 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 CATACAGGAATAAACAATGTGTATATATATACAGGCGAGTATAGACAGGAGTATCC 153

QY 142 AATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGTATTACAAAAGATTTTGGG 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 CATTTTAGCAATTTCTCAGTAAAGAACTAATGTTGATACAAATACAACTAGTAGATAT 213

QY 202 TTAGGCTTGTATAGTCT---GAGATAAATCTCACAGCAATTTTACAGATCATATGAC 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 AAAAAAAGTCGCTCTTCTATCGATCCTAACACTTATTCAAACCTTTCAAGGTCATATACT 273

QY 259 CCTACTTATGCAAGCAGTTTTCAGAGGTTTAGTGGTATCATTTGGATATTATGTTAATGAC 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 GTTCAATTTCAAGATAATGCTGCTAGTTTCAGTGGAGCAATTTGGATATTCTTACCCGAA 333

QY 319 TTT---AGGTAAGATTTGAAGTCTTCTATGAGAAATTTTGAACCTGAAAGACAATGGTAC 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 AGCTAAGACTTGAAGTCTTACGAAAAATTTGATGTCAAAGATCCTAAGAC 393

QY 376 CCGTGAATAGCAAGCTACAAATTTTTCGTTCTGCGAAATGCTCAAAATAGTAT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCACGTAATACGCTACTACTGTT 453

QY 436 AAT-----AGTTTATAGTACTAGAGATAACGGGTTGTTGACAAAGTCTCTTAAT 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 CCTGATGCTCAAAAATATACAGTTATGAAGAAATAATGGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTTGTATGATATTGCTAGTGTAGTATTCCTTTAGCACCTTATATGTGTCT 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 ATCAATGGTTGTATGATCTATCTTTTAAATTTAGTCGTATCACCTTATATATGTGCA 573

QY 547 GGTGTTGGTCAGATATATAAAGTTTGTAGGTATATCATTTGCCCTAAAGTTTCTTATCAA 606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 574 GGTATTTGGTGAAGATTTCAATTGAATTTTGTGATCTTTTGCACATTTAAACCTTGTCTTAT 633

QY 607 GTTAAGTTTGGTCTCAACTACCTCTAAATGTTTAACTATGTTGTTGGTGGGTTAT 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 634 GGAAACTAGGATTAGTTATTACTTCTTTCTTAAGATTAAATGTTGCTGGTGGGTAC 693

QY 667 TACCATAAGGTTGTAGGTGATAGGCATGAGAGATAGAAATAGCTTTACCATCCTACTGCA 726
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 694 TATCATAGATTATAGGGAATAAATTTAAATAATTT---AAATGTTAAACCATGTTGTACA 750

QY 727 TTATCTGACGTTCTTAGAATCTCTCAGCTTCGTCTACTTTTAAATCTGATATTATTTGGT 786
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 CTTGATGAATTTCTTAAAGCAACTTCTGCAAGTACACATTAATGTTGCTTATTTTGGT 810

QY 787 TGGGAGATTGGATTTAGATTTGC 809
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 GGTGAAGCTGGAGTAAAGTTTAC 833
```



US-10-901-774-23  
; Sequence 23, Application US/10901774  
; Publication No. US20040265334A1  
; GENERAL INFORMATION:  
; APPLICANT: RIKIHISA, YASUKO  
; APPLICANT: OHASHI, NORIO  
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA  
; TITLE OF INVENTION: CHAFFEENSIS  
; FILE REFERENCE: 22727-04109  
; CURRENT APPLICATION NUMBER: US/10/901,774  
; CURRENT FILING DATE: 2004-07-29  
; PRIOR FILING DATE: 1999-05-19  
; PRIOR FILING DATE: 1998-09-18  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 23  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Ehrlichia chaffeensis  
US-10-901-774-23

Query Match 20.0%; Score 162.2; DB 8; Length 840;  
Best Local Similarity 53.8%; Pred. No. 9.4e-28;  
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTTTACGCTGAGTACATTTGTTTTCTTATCTCGATGGTCTTTTCTGATGCA 81  
DB 34 GTACTTGATCTCTATATATCATTTCTTATCTTATCTTGAATCCTTTCAGCTATAAATCAAT 93

QY 82 AATTTTTCTGAAGGAGGAGGAGGACTTTATATAGGTAGTCAGTATATAAGTTGGTATCCC 141  
DB 94 CATACAGGAATAACACTAGTGGTATATATATACAGGCGAGTATAGACAGGAGTATCC 153

QY 142 AATTTTGTATTTTTTCAGCTGAAGAAACAAATTCCTGGTATTTACAAAAGAAATTTTGGC 201  
DB 154 CATTTTGTAGCAATTTCTCAGTAAAGAAACAAATTTGTTGATACAATACAACTAGTAGGATAT 213

QY 202 TTAGGCTTCTAAGTCT---GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258  
DB 214 AAAAAAGTGGCTTCTATCGATCTCAACATTTTCAACCTTCAGGTCCATATACT 273

QY 259 CCTACTTATGCAAGCAGTTTTTGCAGGGTTTGTGGTATCATTTGGATATTTATGTAATGAC 318  
DB 274 GTTACATTTCAAGATATGCTGCTAGTTTCAGTGGAGCAATTTGGATTTCTTACCCGAA 333

QY 319 TTT---AGGGTAGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTAC 375  
DB 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAATTTTGTATGTCAAAGATCTCTAAGAC 393

QY 376 CCTGAGAATAGCCAAAGCTACAAATTTTTTGGCTTGTCTCGAAATGCTCAAAATAGTGAT 435  
DB 394 TACTCAGCAAAAGATGCTTTTAGGTTTGTCTCTAGCAAGTAAATAGCTACTACTGT 453

QY 436 AAT-----AAGTTTATAGTACTAGAGATAAAGCGGCTTGTGACAAAGTCTCTTAAT 486  
DB 454 CTTGATGCTCAAAATATACAGTTATGAGAAATAATGGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTTGTATGATATTTGTAGTGGTATGATTTCCCTTTAGCACCTTATATGTTGCT 546  
DB 514 ATCAATGGTGTATGATCTATCTTTTAAATAATTTAGTCGTATCACCTTATATGTGCA 573

QY 547 GGTCTGGTGAGATATATAAAGTTTTTAGGTATATCATTTGCTCAAGTTTTCTTATCAA 606  
DB 574 GGTATTTGGTGAAGATTTCAATGAAATTTTTTGATACCTTTGCACATTAACCTTGTCTATCAA 633

QY 607 GTTAAGTTTGGTGTCAACTACCCCTAAATGTTTAAATACTATGTTTGGTGGGGGTAT 666  
DB 634 GGAATACTAGGTAATGATTACTTCTTCCCTAAGATTAATGATTTGCTGGTGGGTAC 693

QY 667 TACCATAAGGTTGTAGGTGATAGGCAATGAGAGAGTAGAAATAGCTTACCATCTCATGCA 726





327	Qy	AGATTGTAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTACCCCTGAGAAATAG	386
351	Db	 TGAGGTAGAAATGGTCTTATGAAGAATTTGACGTTTAAAAATCCTGGTAAATTTACGTAGTAAAG	410
387	Qy	CCAAAGCTACAAATTTTTTGGCTTCTCGAAATGCTACAAATAGTGTAT-----	435
411	Db	 TGAAGCCTTCAGGTATATTTGCTTTAGCAAGAGGAATTTGATAATCTTCAAAAATATCCTGA	470
436	Qy	----AATAAGTTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAGTCTCTTAATGTAATA	491
471	Db	 AACAAATAAGTATGTTGTTATAAAGAACATGGCTTATCTGTGCGATCCCATTAATATCAAA	530
492	Qy	TGTTGTTTATGATATTTGCTAGTGGTAGTATTTCTTTAGCACCTTATATATGTGCTGGTGT	551
531	Db	 TGGCTGTTATGATTTTCTTTTAAACAAATTTTAAAGATACACCTTACATATCGGTAGGGTT	590
552	Qy	TGGTGACAGATATATAAAAGTTTTTTTAGGTATATCAATGCGCTAAGTTTTCTTATCAAGTTAA	611
591	Db	 TGGTGGGGACATATAGAAATTTTTTAGTGCCTGTAAGTTTTTAAATTTGCTTATCAAGGTAA	650
612	Qy	GTTTGGTGTCAACTACCTCTAAATGTTAAATACTATGTGTTGGTGGGGGTTATTACCA	671
651	Db	 GGTAGGTATCAGTTATCCATTATTTCTCTAAATGATTAATATTTGCTGACGGATTTTACCA	710
672	Qy	TAAGGTTGTAGGTGATAGCGATGAGAGAGTAGTAATAAGCTTACCATCCTACTGCAATTATC	731
711	Db	 TAAGGTCAATAGGAATAAATTTTAAATTTTAAATGTTCAACAGTTGTTAGTCTTAAACAG	770
732	Qy	TGACGTTCTCAGAACTACTTTCAGCTTCTCGTACTTTTAAATACTGATTAATTTTGGTTGGGA	791
771	Db	 TCA---TCCTAGTCTACTTTTGCAGTAGCTACTCTTAAATGTTGAGTATTTTCCGTAGTGA	827
792	Qy	GATTGGATTTAGATTT	807
828	Db	 ATTTGGGTTAAAAATTT	843

```

RESULT 14
US-10-901-774-25
; Sequence 25, Application US/10901774
; Publication No. US20040265334A1
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/901,774
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 25
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
US-10-901-774-25

```

Query Match	18.5%	Score 150.8;	DB 8;	Length 852;
Best Local Similarity	54.6%	Pred. No. 4.5e-25;		
Matches 402;	Conservative	0;	Mismatches 307;	Indels 27;
Gaps	4;			

Qy	96	GAGGAGACCTTTATATATAGTAGTACGATATAAAGTTCGTATTCCTCAATTTTACGTAAATTT	155
Db	111	GTGCACCTGGCTATATGTCAGTGGACAATAAACCCTACTGTTCCTACTTTAGTAAATTT	170
Qy	156	TTCAGCTGAAGAAACCAATTCCTGGTATTACAAAAGAATTTTTCGGTTAGGCTTTGTATTA	215
Db	171	TTCACTTAAAGAAACCTTATCTCGACATTAAGAGTATTATAGGACCTAGCAAAAGATATTTAA	230

216	Qy	GTCT-----GAGATAAATACTCACGCAATTTTACACGATCATATGACCCCTACTTATGC	260
231	Db	GTCTATTACAGATATACAACAATAAAAAAATTCACATTCCTTATTAACACAAATTTCA	290
270	Qy	AAGCAGTTTTGCAGGGTTTAACTGGTATCATTTGGATATATTGTTAATGACATTT---AGGGT	326
291	Db	AGATNATGCTGTTAGCTTCAGTGCAGCTGTTGGATATATTTCCTCAAGACAGTCCAAGGGT	350
327	Qy	AGAATTTGAAGGTTCTTTATGAGAAATTTTGAACCTGAAAGACAAATGGTACCTCGAGNAATAG	386
351	Db	TGAGGTAGAAATGGCTTTATGAAGAAATTTGACGTTTAAAAATCCTCGTAATTACGTAGTAGAAG	410
387	Qy	CCAAAGCTACAAATTTTTTGGCTTGTCTCGAAATGCTACAATAATAGTAGAT-----	435
411	Db	TGAAGCCTTCAGGTATATTGCTTTTACGAAGGGAATTGATTAATCTTCAAAAATATCCTTGA	470
436	Qy	-----AATAAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAGTCTCTTAAATGTA	491
471	Db	AACAAATAAGTATGTTGTTTATAAGAAACAATGGCTTATCTGTGCGATCCAATTATAATCAA	530
492	Qy	TGTTTGGTTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCTGGTGGTGT	551
531	Db	TGGCTGTATTGATTTTTCTTTAAACAATTTAAAAAGTATCACCTTACATATGCGTAGGGTT	590
552	Qy	TGGTGCAGATTATATAAGTTTTTTAGGTATATCATTGCCTAAGTTTTCTTATCAAGTTAA	611
591	Db	TGGTGGGACATATATAGAAATTTTTTAGTGCTGAAGTTTTAAATTTGCTTATCAAGGTAA	650
612	Qy	GTTTGGTGTCAACTACCCTCTAAANGTTTAATACTATATGTTGTTGGTGGGGGTTATTACCA	671
651	Db	GGTAGGTATCAGTTTATCCATTATTCTCTAAATATGATTAATTTGCTGACGGATATTACCA	710
672	Qy	TAAGTTGTAGGTATAGGCATGACAGAGTAGAATACTGTACCATCCTACTGTCATTATC	731
711	Db	TAAGGTATAGGAAATAAATTTACAATTTAAATGTTCCAAACAGTTGTTAGTCTTTAACAG	770
732	Qy	TGACGTTCTCTAGAACTACTTCAGCTCTCTGCTACTTTAAATACTGAATATTTTGGTTGGGA	791
771	Db	TCA---TCCTAAGTCTACTTTTTCAGTAGTACTCTTAATGTTGAGTATTTCCGGTAGTGA	827
792	Qy	GATTGGGAATTAGATTT	807
828	Db	ATTTGGGTTAAAAATTT	843

```

RESULT 15
US-09-811-007-41
; Sequence 41, Application US/09811007
; Publication No. US20030185849A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 41
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-2
US-09-811-007-41

```

Query Match 18.1%; Score 147.2; DB 3; Length 840;  
Best Local Similarity 53.2%; Pred. No. 3.1e-24;  
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;

Qy 1 ATGAATTACAAAGATTGTTGTAGGTGTTACGCTGAGTACATTGTTTTTTTCTTATCT 60  
 |||||  
 Db 1 ATGAATTATAGAAATTTCTAGTAAGAAGCGCTTAATCTCATTAATGTCAATCTTACCA 60  
 Qy 61 GATGGTGCCTTTTCTGA-----TGCAATTTTCTGAAGGAGGAGGACTTTAT 111  
 |||||  
 Db 61 TATCAGTCTTTTTCGAGATCCTGTAGTTTCAAGAACTAATGATAACAAGAAGGCTTCTAC 120  
 Qy 112 ATAGGTAGTCAGTATAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAGRAACA 171  
 |||||  
 Db 121 ATTAGTGCAAGTACAAATCCAAAGTATATCACCTTTTAGAAAAATTTCTCTGCTGAAGAACT 180  
 Qy 172 ATTCCTGGTATTACAAAAAGATTTTTCGTTAGGCTTG-----ATAAGTCTGAG 222  
 |||||  
 Db 181 CCTATTAAATGGACAAATCTCTCACATAAAAAAGTTTCGGACTAAAGAAAGATGGTGAT 240  
 Qy 223 ATAAATCTCACAGCAATTTTACACGATCATATGACCCCTACTTAT-----GCAAGCAGT 276  
 |||||  
 Db 241 ATAACAAAAAAGACGATTTTACAAGAGTAGCTCCAGGCATTGATTTTCAAAATAACTTA 300  
 Qy 277 TTTGCGGGTTAGTGTATCATTTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAA 336  
 |||||  
 Db 301 ATATCAGGATTTTCAGGAAGTATTTGGTTACTCTATGGACGGACCAAGAAATAGAACTTGAA 360  
 Qy 337 GGTTCCTTATCAGAAATTTTGAACCTGAAAGACAATGGTACC-----CTGAGAAATAGCCAA 390  
 |||||  
 Db 361 GCTGCATATCACAATTTATCCAAAAAACACCGATAACAATGATCTGATATGTTGAA 420  
 Qy 391 AGTACAAATTTTGTCTCGAAATGCTACAAATGCTACAAATGATGATTAAGTTTATAGTA 450  
 |||||  
 Db 421 TACTATAACAATTTTGCAATTTCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480  
 Qy 451 CTAGAGAAATACCGCGTTGTTGACAACTCTTTAATGTAAGTGTGTTTATGATATTGCT 510  
 |||||  
 Db 481 CTTAAAAAATGACGGCATACTTTTATGTCAATGGTTAATACTTGTCTATGACATTACA 540  
 Qy 511 AGTGGTAGTATTCCTTTAGCACCTTATATGCTGTGGTGTGGTGCAGATTATATAAAG 570  
 |||||  
 Db 541 GCTGAGGAGTATCTTTTCGTACCATATGCATATGCAGGTATAGGAGCAGATCTTATCACT 600  
 Qy 571 TTTTGTAGTATATCATGCTTAAGTTTCTTATCAAGTTAAGTTTGGTGTCAACTACCT 630  
 |||||  
 Db 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTACCAAGGAAAAATAGGTATTAGTTACCT 660  
 Qy 631 CTAAATGTTAATACTATGTTGTTGTTGGTGGGTTATACCATAAAGTTGTAGTGATAGG 690  
 |||||  
 Db 661 ATCACACCAAGTCTCTGCATTTATTTGGTGGATACTTACCATGGCGTTATTTGGTAATAA 720  
 Qy 691 CATGAGAGGTAGAAATAGCTTACCATCTCTGCAATTTATCTGACGTTCTAGAACTACT 750  
 |||||  
 Db 721 TTTGAGAGATACCTGTA---ATAACTCTGTAGTATTAATGATGCTCTCTCAACACCA 777  
 Qy 751 TCAGCTTCTGCTATTTTAAATACATGATATTTTGGTTGGGAGATTGGATTTAGATT 806  
 |||||  
 Db 778 TCTGCTTCAGTAACTCTTGACGTTGGATCTTTGGCGGAGAAATTGGAAATGAGGTT 833

Search completed: December 28, 2005, 15:20:21  
 Job time : 872 secs